(51) International Patent Classification 6 : C07H 21/00	A2	(11) International Publication Number: WO 99/12950
		(43) International Publication Date: 18 March 1999 (18.03.99)

(21) International Application Number: PCT/GB98/02665

(22) International Filing Date: 4 September 1998 (04.09.98)

(30) Priority Data:
9718863.5   6 September 1997 (06.09.97)   GB

(71) Applicant (for all designated States except US): NATIONAL STARCH AND CHEMICAL INVESTMENT HOLDING CORPORATION [US/US]; Suite 27, 501 Silverside Road, Wilmington, DE 19809 (US).

(72) Inventors; and
(75) Inventors/Applicants (for US only): JOBLING, Stephen, Alan [GB/GB]; 19 Burwell Road, Eaton Socon, St. Neots, Huntingdon, Cambridgeshire PE19 3QQ (GB). SCHWALL, Gerhard, Peter [DE/GB]; 61 The Broad Walk, St. Neots, Huntingdon, Cambridgeshire PE19 2SG (GB). WESTCOTT, Roger, John [GB/GB]; 46 Castle Street, Wellingborough, Northamptonshire NN8 1LW (GB).

(74) Agent: KEITH W. NASH & CO.; 90–92 Regent Street, Cambridge CB2 1DP (GB).

(81) Designated States: AU, CA, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: IMPROVEMENTS IN OR RELATING TO STABILITY OF PLANT STARCHES

RVA profiles

— 1302
········ 1328
— 1352
—·— 1365
—— 1369
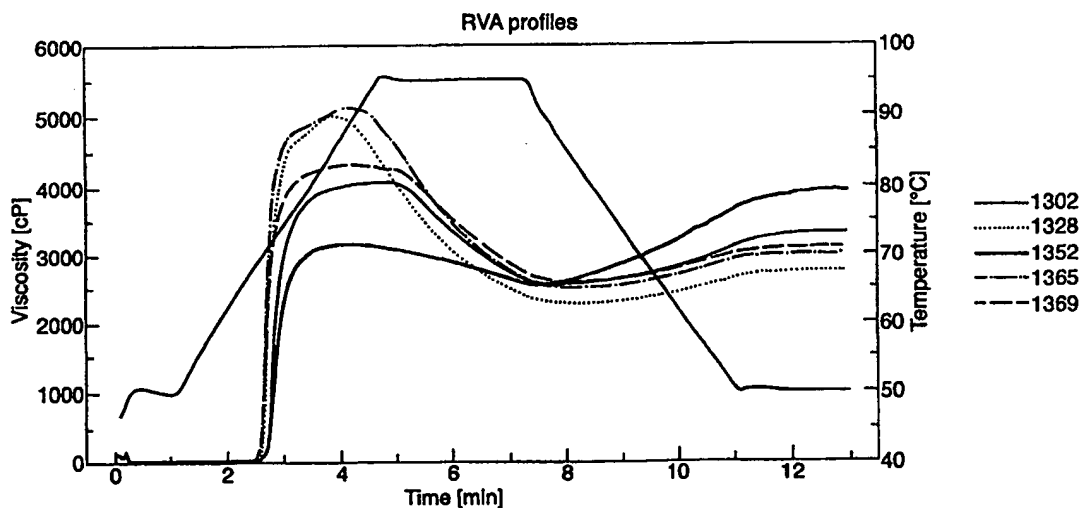
(57) Abstract

Disclosed is a nucleic acid sequence obtainable from potato plants and carrying at least a portion of an isoamylase enzyme gene, together with constructs and host cells comprising the nucleic acid sequence, and methods of using the nucleic acid sequence.

1

Title : Improvements in or Relating to Stability of Plant Starches

Field of the Invention

This invention relates to novel nucleotide sequences, nucleic acid constructs and host cells and organisms comprising said sequences, methods of altering plants using the sequences, and to products obtainable from the altered plants.

Background of the Invention

Starch, the major storage reserve in plants, consists of two main polysaccharides, amylose and amylopectin. Amylose is a linear polymer containing α-1,4 linked glucose units, while amylopectin is a highly branched polymer consisting of a α-1,4 linked glucan backbone with α-1,6 linked glucan branches. In most plant storage reserves, including tubers, amylopectin constitutes about 75% of the starch content. Until recently it was thought that amylopectin was synthesized by the concerted action of two enzymes: starch synthases, which elongate the glucan chains; and starch branching enzymes (SBE), which introduce branches to create amylopectin. It was further considered that it is the chain length specificity of the branching enzymes that determines the starch structure (Tolmasky & Krisman, 1987). Support for this hypothesis was provided by the characterisation of maize (Takeda *et al* 1993) and pea SBE (Burton *et al* 1995).

An alternative hypothesis which is gaining support, suggests that debranching enzymes play an important role in the synthesis of amylopectin and hence starch structure (Ball *et al* 1996; Nakamura,Y. 1996). The evidence for this comes mainly from an analysis of two mutants: *sugary1* mutants of maize synthesise a highly branched water soluble polysaccharide called phytoglycogen and have a defect in a debranching enzyme (Pan and Nelson 1984; James *et al* 1995) and a recently characterised mutant (*STA7*) of *Chlamydomonas*, which has a dramatic effect on starch synthesis, was shown to have a severe reduction in the activity of a debranching enzyme (Mouille *et al* 1996).

2

It is desirable to be able to increase the degree of branching and/or reduce the chain length of the amylopectin in starches, especially waxy potato starches, in order to create more freeze/thaw-stable starches. Very recently, it has been reported that overexpressing the *E. coli* glycogen branching enzyme in a waxy (i.e. amylose-free) potato does lead to an increase in the degree of branching of amylopectin (Kortsee *et al* 1996). However, the increase in branching was only 25%, and the average chain length of the amylopectin was decreased only marginally - by 3 glucose units (from 25 to 22 glucose residues). An alternative approach to obtain a more highly-modified starch would be to decrease the expression of debranching enzymes in the potato tuber using an antisense or sense-suppression approach. To do this efficiently requires the cloning of a cDNA copy of a potato debranching enzyme.

Debranching enzymes (DBE) hydrolyse the α1-6-glucosidic linkages of α-glucan polymers and are classified into two types depending on their substrate specificities: (1) pullulanases (EC 3.2.1.41, which are known in plants as "R- enzymes") which act on amylopectin and pullulan (a linear bacterial polymer consisting of α1-6 linked maltotriose units) but which cannot use glycogen as a substrate; and (2) isoamylases (EC 3.2.1.68) which act on amylopectin and glycogen but cannot use pullulan as a substrate.

Pullulanases have been purified from a number of plants including maize, rice, sorghum, oat, broad bean, potato, sugarbeet and spinach and the corresponding cDNA clones of maize, rice and spinach have been isolated (James *et al* 1995; Nakamura *et al* 1996; Renz *et al* 1995). In potato, both R-enzyme and isoamylase enzyme polypeptides have been characterised and purified (Drummond *et al* 1970; Ishizaki,Y *et al* 1983). A recent publication (WO 95/04826) detailed the purification and peptide sequence determination of a pullulanase type DBE. A cDNA encoding a potato DBE was described in a second patent application (WO 96/19581) which detailed the isolation of a partial length (492 bp) cDNA clone encoding a pullulanase type DBE, as well as the effect of downregulating the expression of this gene in transgenic antisense plants. An unspecified number of lines were generated and some showed a decrease in debranching enzyme activity (not quantitated) which also led to a slight modification in the swelling characteristics of the starch as measured by RVA (Rapid Viscoamylograph). These included a decrease in the

maximum viscosity to about half of the control, and an increase in the setback viscosity (from 3300 cP in the control to 5200 in the most highly-modified plant line). No significant change in the amylose/amylopectin ratio was observed but there was an increase in the phosphate content of the starch (from 9 to 14.6 nmol G6P/mg starch).

The present application relates particularly to the isolation of nucleic acid sequences encoding potato debranching enzymes (especially isoamylase type debranching enzymes) and to the uses thereof.

Summary of the Invention

In a first aspect the invention provides a nucleic acid sequence obtainable from potato plants and carrying at least a portion of an isoamylase-type debranching enzyme (DBE) gene. Those skilled in the art will appreciate that the nucleic acid sequence will comprise at least one region which is characteristic of isoamylase-type debranching enzymes and which is not present in pullulanase type debranching enzymes. The nucleic acid sequence will preferably be at least 200bp long, more preferably at least 300-600bp long. Conveniently the sequence is one which is derived or obtained from potato plants.

The nucleic acid sequence of the invention may be, for instance, a genomic sequence, or may be a cDNA sequence. The nucleic acid sequence of the invention may comprise non-coding portions (which are transcribed, but not translated), and/or may comprise coding portions, which are translated into amino acid residues in the full length native gene. As described below, either translated or non-translated portions may be useful in certain techniques, and both therefore possess utility.

In a particular embodiment, the nucleic acid sequence comprises nucleotides 79-1595 of the nucleic acid sequence shown in Figure 8, or a functional equivalent thereof. (This sequence includes a portion encoding the amino acid sequence MDVVF/YNH, which portion is typically present in the nucleic acid sequences of the invention). Functionally equivalent nucleic acid sequences include sequences which encode substantially the same polypeptide, but which differ in nucleotide sequence from that shown in Figure 8 by virtue of the degeneracy of the genetic code. For example, a nucleic acid sequence may be

4

altered (e.g. "codon optimised") for expression in a particular host cell or host organism, such that the nucleotide sequence differs substantially whilst the amino acid sequence of the encoded polypeptide is unchanged. Typically, functionally equivalent nucleic acid sequences are those which will hybridise under stringent hybridisation conditions (e.g. as described by Sambrook *et al*, Molecular Cloning. A Laboratory Manual, CSH, i.e. washing with 0.1xSSC, 0.5% SDS at 68°C) with the sequence shown in Figure 8. Figures 6A-7B illustrate some related sequences, which are also in accordance with the present invention. These are genomic sequences with large stretches of intron sequence, which introns are not represented in mRNA. The sequences illustrated in Figures 6A-7B thus comprise both coding and non-coding portions.

DNA sequences functionally equivalent to that shown in Figure 8 will preferably comprise at least 200, more preferably 300-600bp, and typically will exhibit at least 85% identity (preferably at least 90%, and more preferably at least 95% identity) with the corresponding region of the DNA sequence shown in Figure 8. Those skilled in the art will readily be able to conduct a sequence alignment between the putative functionally equivalent sequence and that detailed in Figure 8 - the identity of the two sequences is to be compared in those regions which are aligned by standard computer software, which aligns corresponding regions of the sequences.

In a second aspect the invention provides a replicable nucleic acid construct comprising the nucleic acid sequence of the first aspect defined above. The construct may typically comprise a selectable marker (e.g. kanamycin or hygromycin-resistance) and may allow for transcription (and possibly translation) of the nucleic acid sequence of the invention. Conveniently therefore the construct will comprise a promoter (especially a promoter sequence operable in a plant and/or a promoter operable in a bacterial cell), and one or more regulatory signals (such as polyadenylation signals) known to those skilled in the art.

The nucleic acid sequence of the invention can be introduced into plants which express a homologous sequence encoding an isoamylase enzyme, in such a way that the introduced sequence interferes with the expression of the homologous sequence. Thus, for example,

the nucleic acid sequence of the invention may be operably linked in the antisense orientation to a suitable promoter active in the plant into which the sequence is to be introduced, so as to reduce the levels of expression of the homologous isoamylase enzyme present in the plant.

Additionally, it has recently been demonstrated in other experimental systems that "sense suppression" can also occur (i.e. expression of an introduced sequence operably linked in the sense orientation can interfere, by some unknown mechanism, with the expression of the homologous, native gene), as described by Matzke & Matzke (1995 Plant Physiol. *107*, 679-685). Any one of the methods mentioned by Matzke & Matzke could, in theory, be used to affect the expression in a host of a homologous isoamylase gene.

It has been found in some experimental systems that there is no requirement for translation of the introduced sequence in order for an inhibitory effect to be observed, and it is believed that antisense methods are mainly operable by the production of antisense mRNA which hybridises to the sense mRNA, preventing its translation into functional polypeptide, possibly by causing the hybrid RNA to be degraded (e.g. Sheehy *et al*, 1988 PNAS *85*, 8805-8809; Van der Krol *et al*, Mol. Gen. Genet. *220*, 204-212). Sense suppression also requires homology between the introduced sequence and the target gene, but the exact mechanism is unclear. It is apparent that, in relation to both antisense and sense suppression, neither a full length nucleotide sequence, nor a "native" sequence (i.e. the introduced sequence need not be 100% identical to the sequence of the homologous native gene) is essential. Preferably the nucleic acid used will comprise at least 300-600bp of sequence, but by simple trial and error other fragments (smaller or larger) may be found which are functional in altering the characteristics of the plant.

Accordingly, in a third aspect the invention provides a method of altering a plant by inhibiting the expression of an isoamylase gene therein, the method comprising introducing into the plant the nucleic acid sequence of the invention, operably linked (in the sense or anti-sense orientation) to a promoter active in the plant, wherein the isoamylase gene in the plant has homology to the introduced sequence. The introduced nucleic acid will typically comprise at least 200bp, more preferably 300-600bp of sequence having at least

80% (advantageously at least 85%, desirably over 90%) sequence identity with the native isoamylase gene(s) present in the plant.

Advantageously the plant will be a plant which normally comprises large amounts of starch, such as potato, sweet potato, maize, wheat, barley, oat, cassava, pea, rice and the like. It is probable that the greater the homology between the introduced sequence and the native gene, the greater the inhibition of expression of the native gene. Preferably the plant is therefore a potato plant, and/or a plant which has a reduced amylose content (e.g. the "waxy" mutation of potato or the "LAM" [low amylose mutant] strain of pea).

Inhibition of the isoamylase gene might be expected to affect the properties of the starch present in the plant. In particular it might be expected to result in a starch having increased branching and/or shorter chain length. The obtaining of starch having such properties is a particular object of the present invention.

It may also be advantageous to introduce into the plant one or more additional nucleic acid sequences, especially sequences which might also be expected to have an effect on synthesis and/or metabolism of starch in the plant. Particularly preferred are sequences which inhibit the expression of a pullulanase debranching enzyme in the plant. Plant pullulanase sequences are known to those skilled in the art (e.g. WO 96/19581), and a portion of a potato pullulanase cDNA sequence is also disclosed herein. The additional nucleic acid sequence will typically be homologous to a portion of a native gene in the plant and, as described above, may be introduced into the plant operably linked to a suitable promoter in either the sense or the anti-sense orientation.

The introduced sequence of the invention and the one or more additional sequences may be operably linked to a single promoter (which would ensure both sequences were transcribed at essentially the same time), or may be operably linked to separate promoters (which may be necessary for optimal expression). Where separate promoters are employed they may be identical to each other or different. Suitable promoters are well known to those skilled in the art and include both constitutive and inducible types. Examples include the CaMV 35S promoter (e.g. single or tandem repeat) and the GBSS

type I or patatin promoters of potato. Advantageously the promoter will be tissue-specific. Desirably the promoter will cause expression of the operably linked sequence at substantial levels only in the tissue of the plant where starch synthesis and/or starch storage mainly occurs. Thus, for example, where the sequence is introduced into a potato plant, the operably linked promoter may be tuber-specific, such as the patatin promoter.

Those skilled in the art will also appreciate that, if additional nucleic acid sequences are to be introduced into the plant, they may be present on the same nucleic acid plant transformation vector as the portion of isoamylase-encoding sequence, or may be introduced on separate vectors (for example by re-transformation). The sequence of the invention, and the one or more additional sequences if desired, can be introduced into the plant by any one of a number of well-known techniques (e.g. Agrobacterium-mediated transformation, or by "biolistic" methods).

In another aspect, the invention provides a plant cell, or a plant or the progeny thereof, which has been altered by the method defined above. Altered plant cells, into which the sequence of the invention has been introduced, may be grown into plantlets by conventional methods well-known to those skilled in the art. The progeny of an altered plant may be obtained, for example, by vegetative propagation, or by crossing the altered plant and reserving the seed so obtained. The invention also provides parts of the altered plant, such as storage organs. Conveniently, for example, the invention provides tubers comprising altered starch, said tubers being obtained from an altered plant or the progeny thereof.

The invention also provides a method of obtaining starch from an altered plant, the plant being obtained by the method defined above. Starch may be extracted from the plant by any of the known techniques (e.g. milling). The invention further provides starch obtainable from a plant altered by the method defined above, the starch having altered properties compared to starch extracted from an equivalent but unaltered plant. Conveniently the altered starch is obtained from an altered plant selected from the group consisting of cassava, potato, pea, tomato, maize, wheat, barley, oat, sweet potato and rice. Typically the altered starch will have an increased degree of branching and/or

reduced amylopectin chain length, compared to starch extracted from an equivalent but unaltered plant. In general, such altered starch obtained by means of the present invention is found to possess different viscosity charcteristics compared to starch extracted from equivalent, but unaltered, plants. In particular, the inventors have found that starch obtained from plants altered by the method of the invention may have reduced peak viscosity compared to starch extracted from an equivalent but unaltered plant, as determined by analysis of a 7.4% (w/v) starch suspension using a Rapid Visco Analyser series 4 instrument, (available from Newport Scientific, Sydney, Australia) according to the manufacturer's standard stirring and heating protocol. Alternatively, or additionally, the altered starch may have higher setback viscosity compared to starch extracted from an equivalent but unaltered plant, (as determined, for example, by analysis of a 7.4% (w/v) starch suspension using a Rapid Visco Analyser series 4 instrument). Further, the altered starch may have an increased viscosity onset temperature compared to starch extracted from an equivalent but unaltered plant, (as determined by analysis of a 7.4% (w/v) starch suspension using a Rapid Visco Analyser series 4 instrument).

The terms "peak viscosity", "setback viscosity" and "viscosity onset temperature" are well known and understood by the person skilled in the art. By way of explanation for the novice reader, the viscosity of starch preparations is generally investigated by heating an aqueous sample of the starch under controlled conditions of temperature increase, temperature maintenance, and subsequent temperature decrease and with controlled stirring, using an instrument known as a Visco analyser, which produces a viscosity profile called a "viscoamylograph". The "peak viscosity" is the maximum viscosity which is attained during the temperature increase phase. The "setback viscosity" is the viscosity obtained at the end of the viscoamylograph. The viscosity onset temperature is the temperature at which a sudden increase of viscosity is apparent on the viscoamylograph.

For the purposes of the present invention, an "equivalent, but unaltered" plant may be considered as a plant which has a substantially identical genotype to a plant altered by the method of the invention, except for the introduced sequences present in the altered plant of the invention. Also to be considered as an equivalent but unaltered plant are those plants which comprise the introduced sequence but wherein the introduced sequence has

not affected the starch properties of the plant.

In a further aspect, the invention provides a nucleic acid sequence encoding at least a portion of a potato pullulanase type debranching enzyme, the sequence comprising substantially nucleotides 46-774 of the nucleic acid sequence shown in Figure 10, or a functional equivalent thereof. The term "functional equivalent" should be interpreted as explained previously, but with reference (in this aspect of the invention) to Figure 10 rather than Figure 8.

Similarly, the invention provides nucleic acid constructs, and altered host plant cells and organisms, comprising the pullulanase-type DBE nucleic acid sequence defined immediately above.

The invention will now be further described by way of illustrative example and with reference to the accompanying drawings, in which:

Figure 1 shows a comparison of the known amino acid sequences of various putative isoamylase proteins. The shaded area indicates residues (single letter amino acid code) that are identical to the maize protein. Dashes indicate gaps introduced to maximise the alignment. Although the whole sequences were compared, only the most conserved region is shown. Sequences are identified at the right; A10906_1 is an artificial sequence, A37035 is from a *Pseudomonas* species SMP1, HIU32761 is from *Haemophilus influenzae*, ISOA_PSESP is from *Pseudomonas amylodermosa* and ZMU18908_1 is from *Zea mays* (corn);

Figures 2A and 2B show a comparison of the 5' and 3' ends (respectively) of two partial pullulanase type DBE genomic clones recovered from potato by the present inventors (nucleotides that are identical are shown between the two sequences and dashes indicate gaps introduced to optimise the alignment);

Figures 3A and 3B show the DNA sequence (SEQ. ID. NO.s 1 and 4 respectively) and predicted intron/exon structure, including predicted amino acid sequences (SEQ. ID. NO.s

2 and 3) at the 5' and 3' ends respectively of a partial pullulanase type DBE genomic clone (certain restriction sites are also marked);

Figures 4A and 4B show the DNA sequence (SEQ. ID. NO.s 5 and 8 respectively) and predicted intron/exon structure, including predicted amino acid sequences (SEQ.ID. NO.s 6 and 7) at the 5' and 3' ends respectively of a partial pullulanase type DBE genomic clone (certain restriction sites are also marked);

Figures 5A and 5B show a comparison of the 5' and 3' ends (respectively) of two partial isoamylase type DBE genomic clones recovered from potato by the present inventors (nucleotides that are identical are shown between the two sequences and dashes indicate gaps introduced to optimise the alignment);

Figures 6A and 6B show the DNA sequence (SEQ. ID. NO.s 9 and 11 respectively) and predicted intron/exon structure, including predicted amino acid sequence (SEQ. ID. NO. 10) at the 5' and 3' ends respectively of a partial isoamylase type DBE genomic clone (certain restriction sites are also marked);

Figures 7A and 7B show the DNA sequence (SEQ. ID. NO.s 12 and 13 respectively) and predicted intron/exon structure at the 5' and 3' ends respectively of a partial isoamylase type DBE genomic clone (certain restriction sites are also marked);

Figure 8 shows the DNA sequence (SEQ. ID. NO. 14), restriction map and predicted open reading frame (including predicted amino acid sequence, SEQ. ID. NO. 15) of a partial length potato debranching enzyme (isoamylase) cDNA clone (pSJ132, described below); nucleotides 79-1595, including the poly(A) tail, are isoamylase, the other sequence is derived from pT7Blue vector;

Figure 9 is a schematic representation of the plant transformation vector pSJ138;

Figure 10 shows the DNA sequence (SEQ. ID. NO. 16), restriction map and predicted open reading frame (including predicted amino acid sequence, SEQ. ID. NO. 17) of a

pullulanase type partial length potato debranching enzyme cDNA clone (pSJ98, described below); nucleotides 46-774 are DBE, the other sequence is derived from pT7Blue vector;

Figure 11 is a schematic representation of a "double" plant transformation vector, comprising partial isoamylase- and pullulanase-type debranching enzyme sequences operably linked in the anti-sense orientation to separate GBSS promoters;

Figure 12 shows a zymogram (using a 5% polyacrylamide gel containing 1% amylopectin) of tuber extracts from potato lines transformed with an antisense isoamylase construct in accordance with the invention; and

Figure 13 shows viscosity (RVA) profiles of starch from potato lines transformed with an antisense isoamylase construct, and from a wild type control (1302).

EXAMPLES

Example 1: **Isolation of a partial cDNA clone of a potato isoamylase.**
Throughout the examples described in the present specification, standard procedures were performed essentially according to Sambrook *et al* (1989). DNA sequencing was performed on an ABI automated DNA sequencer and sequences manipulated using DNASTAR software for the Macintosh computer.

Oligonucleotides
The following oligonucleotides (SEQ ID. NO.s 18-29 respectively) were employed in the examples described below:

RoRidT$_{17}$ AAG GAT CCG TCG ACA TCG ATA ATA CGA CTC ACT ATA GGGA(T17)
Ro   AAG GAT CCG TCG ACA TC
Ri   GAC ATC GAT AAT ACG AC
PDBE1   GAC/T GAC/T GTn AAA/G CCT GAA/G GG
PDBE2   ATC AAA TCT/G A/GAA nCC ATC

12

DBE1	ATG GAT GTT GT(T/C/A) TA/TC/T AAT/C CAT
DBE2	TCA AAA/T CG/TA/G AAi CCA TC
PIA1	GAG GCA TTG ACA ACA GTG TG
PIA2	CGT GAC ATG GTA CGG CAG TT
PIA3	AAC TGC CGT ACC ATG TCA CG
PIA4	GGT TAT AGA TGG CAG CCG TT
PIA5	AAC GGC TGC CAT CTA TAA CC

Potato DNA isolation

Genomic DNA was isolated from potato leaves using a modified CTAB (cetyltrimethyl ammonium bromide) method (Murray and Thompson, 1980). Leaf tissue (4 gm) was frozen in liquid nitrogen and ground in a mortar and pestle and transferred to a 50 ml polypropylene centrifuge tube (Falcon) to which was added 40 ml of extraction buffer (2% w/v CTAB, 100 mM Tris-HCl pH 8.0, 1.4 M NaCl, 10 mM β-mercaptoethanol and 20 mM EDTA) and incubated at 65°C for 30 min. After cooling, the mixture was split into two tubes and 15 ml of chloroform was added to each followed by gentle (end over end) mixing for 15 min at room temperature. The tubes were then centrifuged for 5 min at 5000g and the pooled aqueous phases were re-extracted with 5 ml of chloroform. The DNA was then precipitated by incubating on ice for 30 minutes with an equal volume of isopropanol and centrifuged for 5 minutes at 3000g. The pellet was washed once with 70% ethanol, briefly air-dried and resuspended in TE pH 8.0 containing 20 μg/ml RNase A. After incubation at room temperature for 30 minutes the DNA was extracted with 5ml of chloroform, one tenth volume of 3 M sodium acetate (NaAc) pH 5.2 was added and the DNA precipitated with two volumes of ethanol. After centrifugation and washing as above, the DNA was resuspended in 2 ml of TE buffer at pH 8.0 and estimated to be approximately 150 ng/μl by fluorometric measurement.

RNA isolation and cDNA synthesis

RNA was isolated from potato plants by the method of Logemann et al (1987). Tuber material consisted of small tubers (<1 gm) harvested from a mature plant. Leaf material was from source leaves (>10 cm) of a healthy vigorously growing plant from the greenhouse. Five microgram of RNA was heated in water for 3 minutes at 65°C and then

13

reverse transcribed in a 50 μl reaction using 10 pmol of RoRidT17 as primer and MMLV reverse transcriptase (Stratascript 50 units; reaction conditions 50mM Tris-HCl (pH 8. 3), 75mM KCl, 10mM DTT, 3mM MgCl₂, 0.5mM each dNTP, 1U/μl RNasin and incubation for 1 hour at 37°C). The reaction was then diluted to 200 μl with TE pH 8.0 and stored at 4°C.

Isoamylase cloning

No amino acid sequence information for isoamylase enzymes from potato was publicly available at the time this work was performed, therefore the cloning strategy relied only on conserved sequence regions among known or putative isoamylase type debranching enzymes from other sources. As shown in Figure 1, there are two regions within isoamylase proteins that are conserved. One of these is the same DGFRFD sequence found in the pullulanase enzymes (see example 3) and the second (MDVVF/YNH) is found approx 70 aa from this sequence towards the N-terminus. Degenerate oligonucleotide primers were made corresponding to these sequences (DBE2 & DBE1 respectively, see above), taking into consideration the codon usage in plant genes in order to reduce the degeneracy of the primers. These were used in a PCR reaction using first strand cDNA from leaf or tuber RNA or genomic DNA as template.

Approximately 50ng of potato genomic DNA or 5μl of first strand cDNA were used as template in a 50μl PCR reaction containing 25pmol of each degenerate primer (DBE1 & 2) under standard PCR conditions (30 cycles of 94°C 60 seconds, 50°C 60 seconds, 72°C 60 seconds). The PCR products were separated on a 1% agarose gel in TAE buffer. The expected size fragment (240 bp) was amplified from both leaf and tuber cDNA templates although the signal was stronger in the latter. However when genomic DNA was used as template, 5 strong bands of approx 2.4, 2.0, 1.8, 1.6 and 0.6 kb were amplified indicating that intron(s) are present within this portion of the genes. This pattern was very reproducible and only occurred when both primers were used in the PCR reaction. These 5 bands were cut out and purified using Qiaquick columns (Qiagen) and then cloned in the pT7Blue vector (Novagen).

Colonies were screened by PCR and positives were grown up and DNA isolated from a 50

14

ml culture by alkaline lysis and PEG precipitation. One clone of each of the 5 PCR fragments was sequenced from either end and in one direction only, therefore the sequence cannot be guaranteed 100 % correct.

The exon sequences of the 1.8 and 1.6 kb bands (clones #43 & #34, designated pSJ134 and pSJ135 respectively) were found to be identical to pSJ98 (a pullulanase type DBE partial clone, described below) indicating that they are pullulanase type debranching enzymes. The two clones pSJ134 and pSJ135 are not identical, sharing 96.3% identity over the entire sequence at the 5' end, and 89.8% identity over 160 nucleotides at the 3' end, after which the two sequences diverge. Figure 2A shows a comparison of the sequences of clones 134 and 135 at their 5' ends, whilst Figure 2B shows a comparison of their sequences at the 3' end. Figures 3A and 3B illustrate the presumed intron/exon structure of the 5' and 3' ends (respectively) of clone 134. Similarly, Figures 4A and 4B illustrate the presumed intron/exon structure of the 5' and 3' ends respectively of clone 135. In both clones, at the 5' end two exons can be distinguished, separated by a 101 bp intron and the remainder of the exon sequences are present in the 3' end.

The 2.4 and 2.0 kb bands (clones #44 & #42, designated pSJ133 and pSJ137 respectively) showed most homology to the isoamylase genes (data not shown). Figure 5A shows a comparison of the 5' ends of these clones, whilst Figure 5B compares their 3' ends. As with the pullulanase clones, clones 133 and 137 were not identical (96 and 97% identity at the 5' and 3' ends respectively). Figures 6A and 6B illustrate the presumed intron/exon structure of the 5' and 3' ends (respectively) of clone 133. Similarly, Figures 7A and 7B illustrate the presumed intron/exon structure of the 5' and 3' ends respectively of clone 137. Only one intron can be distinguished at each end.

The 0.8 kb clone, although it had the correct primers on both ends was not a debranching enzyme but had homology to a rice Expressed Sequence Tag (RICR2232A) which encodes a 70kDa endomembrane protein.

A new sense primer (PIA1) was designed to the isoamylase sequences just downstream of the DBE1 primer and this was use in a 3'RACE PCR to clone the 3' end of the isoamylase

15

gene. Five μl of first strand cDNA and 25 pmol of each primer (PIA1 and Ro) were used in a 50 μl PCR reaction (94°C 2 min followed by 30 cycles of 94°C 45 seconds, 55°C 25 seconds, 72°C 2 min). One μl of this reaction was then used in a second PCR amplification using PIA1 and R_i primers. A major band of approx 1.5 kb (the expected size) was amplified from both leaf and tuber cDNA templates. The tuber PCR fragment was purified on a 1% TAE agarose gel and then cloned in pT7Blue (clone designated pSJ132). The sequence and predicted ORF of this partial length clone is shown in Figure 8. The ORF contains 429 aa and is 77.0% identical to the maize isoamylase protein encoded by the sugary-1 gene (data not shown).

Example 2: **Transformation of plants with antisense isoamylase**

A plant transformation vector with the hygromycin selectable marker was constructed in which the 1.5 kb isoamylase cDNA from plasmid pSJ132 was inserted in an antisense orientation under the control of the granule bound starch synthase (GBSS) promoter. This vector (pSJ138) was then used to transform wild type and waxy potato plants.

Construction of plant transformation vector

A plant transformation vector (pSJ138) was constructed by inserting the 1.5 kb isoamylase cDNA from plasmid pSJ132 (as an *Xba* I-*Sac* I fragment) in an antisense orientation under the control of the granule bound starch synthase (GBSS) promoter into the plant transformation vector pSJ39 which had been cut with the same enzymes. pSJ39 is a modified GPTV-HYG vector (Becker *et al* 1992) containing a 0.8 kb GBSS promoter and was constructed as follows: first the *Bam*HI site between the hygromycin selectable marker and the gene 7 polyadenylation signal in pGPTV-HYG was destroyed by cutting and filling in with klenow polymerase to create pSJ35. The *Hind*III-*Eco*RI fragment of this vector (containing the GUS gene & nos polyadenylation signal) was replaced with the *Hind*III-*Eco*RI fragment (containing the GBSS promoter-GUS-nos poly(A) cassette) from plasmid pGB121 (obtained from R Visser, Wageningen) to create plasmid pSJ39. Plasmid pGB121 was constructed by inserting the 0.8 kb GBSS promoter from genomic clone LGBSSwt-6 (Visser *et al* 1989) as an *Hind*III-*Nsi*I (klenow filled in) fragment into the *Hind*III-*Pst*I(T4 DNA polymerase blunted) sites of plasmid pBI121 (Jefferson *et al* 1987). Construction of pSJ138 is illustrated schematically in Figure 9.

16

In Figure 9 the black line represents the DNA sequence. The hashed line represents the bacterial plasmid backbone (containing the origin of replication and bacterial selection marker) and is not shown in full. The filled triangles represent the T-DNA borders (RB = right border and LB = left border). Relevant restriction enzyme sites are shown above the black line with the approximate distances (in kilobases kb) between sites marked by an asterisk shown underneath. The thinnest arrows represent polyadenylation signals (pA-nos = nopaline synthase, pAg7 = Agrobacterium gene 7), the intermediate arrows denote protein coding regions (HYG = hyromcin resistance gene etc) and the thick arrows indicate promoter regions (P-GBSS = granule bound starch synthase and P-nos = nopaline synthase).

Potato transformation protocol

Transformed potato plants cv Desirée were produced by co-cultivation of explants with Agrobacterium tumefaciens LBA 4404 containing plant transformation vector plasmids. Two explant types were used, microtubers and leaf fragments, both produced from stock cultures, requiring different methods in culture. Stocks of plantlets were maintained of wild type untransformed Desirée or selected lines previously transformed with antisense genes coding for enzymes of the starch biosynthetic pathway by regular transfer of single node explants using Murashige and Skoog media (1962 Physiol. Plant 15, 473) solidified with 0.8% agar with the sucrose concentration reduced from 3% to 1%. Microtuber explants were produced by transferring nodes to the same medium with the sucrose concentration increased to 8% plus the addition of 2.5mg/l benzylaminopurine. All cultures were grown at 22°C with illumination from fluorescent tubes at 40µE/sqm for 16 hours, except those cltures grown to produce microtubers, which were kept in darkness. The protocol for microtuber explants, and that for leaf explants (based ona modification of that published by Rocha-Sosa et al 1989, EMBO J. 8, 23), are detailed below.

Explants were initially precultured on the appropriate multiplication medium for two days before co-cultivation. Co-cultivation was for 10 minutes in an overnight culture of the Agrobacterium before explants were blotted on filter paper to remove excess bacteria. After blotting explants were transferred to filter paper on feeder layers and incubated in darkness at 22°C for two days. The feeder layers consisted of a 9cm petri dish containing

the appropriate multiplication media covered by 2ml of stationary phase tobacco suspension culture cells which were in turn covered by two layers filter paper.

After two days the explants were removed and washed in Murashige and Skoog medium (without sucrose or agar) containing 500μg/ml cefataxime, blotted on filter paper and transferred to multiplication medium containing 500μg/ml cefataxime. After a further five days the explants were further transferred to the same medium containing the selection agent. Selection was effected by the use of the NPTII gene in the construct and addition of kanamycin in the tissue culture process or the HPTII gene with additions of hygromycin. To achieve double transformations lines resistant to kanamycin containing the NPTII gene and a gene to modulate starch biosynthesis were subsequently transformed with a construct containing the HPTII using hygromycin selection. This was reversed if the line used had been originally transformed with a HPTII containing construct and NPTII and kanamycin were used for selection.

Tissue culture and selection using microtuber explants was performed as follows. Microtubers were taken from stock when they were sufficiently large and mature to use. This occurred after 6-8 weeks from transfer of nodes to the high sucrose medium when they weighed more than 30mg and had started to turn from white to pink. Stocks were used for several months after tuber initiation provided they remained dormant and no new shoots or stolons had emerged from the bud end. Each tuber was cut in half longitudinally through the terminal bud and the cut end put in contact with the medium. The multiplication medium (Z5) used contained the salts and vitamins of Murashige and Skoog plus 3% sucrose, 0.2mg/l indole acetic acid, 5 mg/l zeatin and 0.8% agar. Kanamycin was added at 100mg/l when NPTII was used and hygromycin at 15mg/l when HPTII was used. The explants were transferred to fresh medium every two weeks. Explants had two transfers on medium Z5, followed by one transfer on the same medium, except indole acetic acid and zeatin were replaced by 10mg/l of gibberellic acid (medium MSG). At each transfer all shoots were removed and discarded. The final transfer was to hormone-free Murashige and Skoog medium with the sucrose concentration reduced from 3% to 1%. Shoots were allowed to develop at this stage.

Tissue culture and selection using leaf explants was performed as follows. Expanded leaves were taken from stock cultures and cut into fragments comprising a half or a third of the leaf. The multiplication medium used contained the salts and vitamins of Murashige and Skoog medium plus 1.6% glucose, 0.02 mg/l naphthyl acetic acid, 2mg/l zeatin riboside and 0.8% agar. Kanamycin was added at 100mg/l when NPTII was used and hygromycin at 5mg/l when HPTII was used. Cultures were transferred every two weeks to the same medium. After four transfers the cultures which were vitrified or showed poor shoot development were transferred to liquid Murashige and Skoog medium containing 1.6% sucrose only.

Transformation was confirmed as follows. Shoots which developed from either tuber of leaf explants after the multiplication/organogenesis phase were transferred individually to hormone-free Murashige and Skoog medium containing 1% sucrose and 100 mg/l kanamycin or 15mg/l hygromycin depending on the selection which was used. Only one shoot was taken from each explant unless the origin of separate shoots were well separated on the explant. Shoots which rooted on selection within two weeks were transferred to the same medium without selection and given an identifying number. When growth allowed sufficient tissue to be sampled, DNA was extracted from a few mgs of leaf and stem tissue of the culture using the method published by Edwards *et al* (1991 Nucleic Acids Research 19,(6) 1349) and transformation was confirmed by PCR using gene specific primers. Material was left in culture to allow individual shoots to be micropropagated. Positive shoots were micropropagated to be stored *in vitro* and also to provide planting material.

Growth of Transformed Lines
Five plants from each line, when 50-100mm high, were transferred to compost made up of 50% horticultural sand and 50% Levington F2 peat-based compost in 20mm square modular pots and maintained at 20°C (day) and 15°C (night) in a growth room illuminated for 16 hours with high pressure sodium lamps at an illumination of 400 µE/sq m. After watering they were covered within a small propagator and shaded from illumination. After 7-10 days when new growth was apparent the shading was removed and the plants grown on, with an intermediate potting to 70mm diameter pots. Finally the five plants were grown to maturity together in a 250mm diameter pot in Levington F2 compost under glass.

Sixteen to eighteen weeks after transfer from culture when the foliage began to die down all the tubers were harvested.

Representative samples of the tubers were stored for analysis and possible regrowth and 150-200g fresh weight was taken for starch extraction.

Starch extraction

Washed tubers were diced and homogenised to a juice by passage through a Braun MP 75 centrifugal juicer. The juice was diluted with tap water to a volume of approximately 3L in a 5L conical flask. After 30 minutes the unsedimented supernatant liquid was discarded and the remaining solids washed again with a further 3L of water. After a further 30 minutes the liquid was decanted and discarded. The slurry of solids was poured through a 500µm metal mesh filter which retained the bulk of the cell debris. The crude starch which passed through the filter was washed in the same manner until no protein foam remained and only white starch settled on the base of the flask below a clear supernatant liquid. The starch was harvested on filter paper using a Buchner funnel and washed on the funnel with two 500ml volumes of distiled water. Finally the starch was washed with 500ml of acetone and dried in a fume hood for approximately 30 minutes.

Starch analysis

The degree of branching of the starches was determined by ^{1}H NMR spectroscopy (Gidley 1985), and more detailed analysis performed as dscribed in example 5 below.

Example 3: **Cloning of a pullulanase type debranching enzyme from potato.**

A PCR based cloning strategy was devised for isolating pullulanase type debranching enzymes from potato, using conserved domains within the known cloned gene sequences, of which spinach was the only plant gene (genbank SOPULSPO_1). The amino acid sequence DGFRFDLMG was almost completely conserved (data not shown) and this is found at aa 546-554 of the spinach protein (964 aa in total). It was hoped therefore that this would also be present in the potato gene, as the sequence DGFRFD is also completely conserved in plant starch branching enzyme genes which belong to the same superfamily of amylolytic genes.

20

Peptide sequences from the purified potato debranching enzyme which were published in WO 95/04826 were also compared to the spinach sequence; of the 12 peptide sequences 10 made possible matches to the spinach sequence and of these peptide number 1 (RTLLVNLDSDDVKPEGGDNL, corresponding to aa 300-319 of the spinach sequence, identical residues underlined) was chosen to make a degenerate oligonucleotide sense primer (PDBE1;20mer,64x) to be used in PCR with a degenerate antisense primer designed to the sequence DGFRFD (PDBE2; 18mer, 16x). These primers were used in a PCR reaction with tuber or leaf first strand cDNA as template.

Five μl of cDNA was used as template in a 50 μl PCR reaction with 50 pmol of each degenerate primer (PDBE1 & 2) under standard PCR conditions (28 cycles of 94°C 45 seconds, 48°C 25 seconds, 72°C 90 seconds). The PCR products were separated on a 1% agarose gel in TAE buffer. A strong band of the expected size (672bp) was amplified from the tuber cDNA and a much weaker band from the leaf cDNA. The tuber amplified band was cut out and purified by binding to DEAE paper and then cloned in the pT7Blue vector (Novagen). Colonies were screened by PCR and positives were grown up and DNA isolated from a 50 ml culture by alkaline lysis and PEG precipitation. The resulting plasmid was designated pSJ98 and the sequence is shown in Figure 10. Comparison of this sequence to the spinach gene showed it to be 75.7% identical at the DNA level and 75.5 % at the protein (data not shown). It is concluded therefore that this gene belongs to the pullulanase class of debranching enzymes. In all, three clones of the potato DBE PCR in pT7Blue were sequenced (#4, 5 & 6) #4 = pSJ98, #5 = pSJ136 opposite orientation (17305-506). Clone #6 was found to have a 34bp deletion approx 100 bp from the 5' end, which caused a shift in the reading frame resulting in a truncated polypeptide.

Example 4 Construction of double antisense isoamylase and pullulanase plant transformation vector

A cassette consisting of the 0.7 kb pullulanase type DBE cDNA from pSJ98 in an antisense orientation between the GBSS promoter and CaMV poly A signal (Guerineau et al 1988), was inserted as a blunt fragment into the unique Pme I site of the plant transformation vector pSJ138. A diagram of the resulting vector is shown in Figure 11. In Figure 11, the symbols used have the same meanings as in Figure 9: pA-CaMV is the CaMV

21

polyadenylation signal. .. This vector may then be used to transform potato plants, substantially as described in Example 2.

Example 5 Analysis of potato plants transformed with Isoamylase antisense construct (from example 2)

Zymograms

There is no specific assay for debranching enzymes. Therefore zymograms are usually used to determine the presence or absence of debranching enzymes. Protein extracts were separated on native 5% polyacrylamide gels containing 1% potato amylopectin. After separation the gel was incubated for 1-3 hours in 100 mM citrate buffer pH 5.5 containing 10 mM EDTA, 1 mM acarbose, and 2.5 mM DTT. The gel was then rinsed twice in deionized water and stained with iodine (Lugol's solution). Debranching enzymes stain as blue bands against a purple background, whereas amylases appear as clear bands.

In order to distinguish between the two classes of debranching enzymes, isoamylases and pullulanases, a second set of protein extracts was separated on another polyacrylamide gel containing 1% amylopectin. After the run the proteins were transferred onto a 10% polyacrylamide gel containing 1% red pullulan (Megazyme, Wicklow, Ireland) using standard Western blotting apparatus. The red pullulan containing gel was then incubated as described above for the amylopectin containing gel. Pullulanase bands appear as clear bands against the red background. The debranching enzyme bands in the amylopectin gel can thus be classified into isoamylase bands (which do not appear in the red pullulan containing gel) and pullulanase bands (which give clear bands in red pullulan containing gels).

In Figure 12 tuber extracts of ten different antisense lines were analysed in a zymogram as described above. The upper band with a low electrophoretic mobility is a debranching enzyme as it stains blue in the amylopectin containing gel against a purple background. By comparison with the band pattern of a red pullulan containing gel (data not shown) it could be demonstrated that this band is not a pullulanase and can therefore be classified as an isoamylase. Transformed potato line 1352 clearly lacks this isoamylase band. In this line

isoamylase expression was successfully inhibited by the antisense construct. The lower band apparent in Figure 12 is an amylase which stains as a clear zone in a purple background. Depending on incubation time, two additional debranching enzyme bands could be detected which are of the pullulanase type and not affected in line 1352.

Analysis of starch by RVA (Rapid Visco Analyser)

In addition to the examination of isoamylase enzyme expression by zymograms, the physical properties of tuber starch was analysed, in order to test whether the lack of the isoamylase enzyme alters the starch properties.

Viscosity development of starches was analysed using a Rapid Visco Analyser Series 4 instrument (Newport Scientific, Sydney, Australia). For the 13 minute profile, 2 g of starch was weighed into a sample cup and 25 ml of water was added to give a final starch concentration of 7.4 % (w/v) and the analysis was performed using the standard stirring and heating protocol according to the manufacturer's instructions. Viscosity was measured in centipoise (cP).

Figure 13 shows the viscosity profiles of tuber starches from plants transformed with the antisense isoamylase construct. The linear plot shows the temperature profile. In addition a wild type control (line 1302, unbroken thin line) is shown. Starch from line 1352 (solid thick curve) shows a clearly reduced peak viscosity and a significantly higher setback viscosity compared to the wild type control (1302) and starches of transformed plants with no change in isoamylase expression level (lines 1328 - thin dotted curve; 1365 - thin dotted/dashed line; and 1369 - thin dashed line). In addition the viscosity onset temperature (pasting temperature) is slightly increased for line 1352. These data clearly indicate that the reduction of isoamylase activity (in line 1352) changes the starch structure and composition resulting in novel physico-chemical properties.

Discussi n

When this work was started, no cDNA clones of potato debranching enzymes had been reported although some peptide sequences from a purified potato pullulanase were published in patent WO 95/04826. The present inventors were able to clone a partial

length (672 bp) cDNA .clone of a potato pullulanase (pSJ98, Figure 10). A second publication (WO 96/19581) by the authors of WO 95/04826 then described the cloning of a full length cDNA of a spinach DBE (pullulanase type) as well as a PCR fragment of a potato pullulanase. The latter PCR fragment (492 bp) was isolated using primers derived from the sequence of a genomic clone which hybridised to a degenerate oligonucleotide very similar to the PDBE1 oligo which was used in the experiments described here. However the deduced amino acid sequence of the PCR fragment did not correspond exactly to the two peptide sequences from the purified protein (RTLLVNLDSDDVKPEG$\underline{G}$DNL, SEQ. ID. NO. 30; and RLSSAGITHVHL$\underline{L}$PTYNFA, SEQ. ID. NO. 31; differences are shown underlined). The authors of WO 96/19581 concluded that this was a result of the protein being purified from the variety Desirée whereas the cDNA library was constructed from the variety Berolina. However, the peptide sequences also do not match exactly the sequence of pSJ98 (which was isolated from Desirée) suggesting that more than one pullulanase type DBE gene exists in potato. This is supported by the comparison the of the two PCR fragments which show only 93.6% identity at the DNA and 90.7% at the protein level with one amino acid insertion in the prior art clone (data not shown).

Further evidence that more than one pullulanase gene exists was provided by the cloning of two PCR fragments (1.8 and 1.6 kb) from genomic DNA (pSJ134 & 135). Sequence obtained from the ends of the clones indicated that they differed in the size of one intron (compare Figures 3A/B and 4A/B). The exon sequence of clone pSJ134 was identical to pSJ98 at the amino acid level but those of pSJ135 differed at one amino acid (results not shown).

The inventors have also cloned two PCR fragments (2.4 and 2.0 kb; pSJ133 and 137 respectively) from genomic DNA of potato which have significant homology to isoamylase genes. These clones differ in the size, and perhaps number, of introns and are about 96-97% identical (Figure 4), suggesting that they are derived from different genes. A cDNA fragment of approx 1500 bp was obtained by 3'RACE; this cDNA encodes 429 aa which are 77% identical to the maize isoamylase gene and the inventors are confident that this gene is an isoamylase type DBE.

24

It is possible that the two different isoamylase-type DBE genomic sequences isolated by the inventors relate to two different genes having slightly different functions or roles in the potato plant, and that similar multiple isoamylase-type genes exist in other plants. Different isoamylase-type genes might, for instance, be subjected to different regulatory controls. As the genomic sequences recovered are not full length genes, it is possible that the respective gene sequences may diverge in the uncloned portions. It may therefore be desirable to ensure that, in the various methods of the invention defined previously, a single sequence is used which is sufficiently long, and has sufficient homology to both isoamylase-type sequences that expression of both native isoamylase-type genes (if present) is inhibited, or, in the alternative, to ensure that two different sequences are employed to inhibit expression of the respective native isoamylase-type genes. The same comments apply equally to the pullulanase-type sequences discovered by the inventors.

Both pullulanase and isoamylase genes appear to be expressed in tubers and leaves since PCR bands of the correct size can be amplified from either tuber or leaf first strand cDNA templates. The intensity of the amplified band is stronger in tubers for both DBE, possibly indicating that the expression level of both genes is higher in tubers, but this should be confirmed by northern analysis.

References

Ball et al., (1996) Cell 86, 349-352.

Becker et al., (1992) Plant Molecular Biology 20, 1195-1197.

Burton et al., (1995) Plant J. 7, 3-15.

Drummond et al., (1970) FEBS Lett 9, 136-140.

Frohman et al., (1988) Proc. Natl. Acad. Sci. USA 85, 8998-9002.

Gidley (1985) Carbohydrate Research 139, 85-93.

Guerineau et al., (1988) Nucl. Acids Res. 16, 11380.

Ishizaki et al., (1983) Agric. Biol. Chem. 47, 771-779.

James et al., (1995) Plant Cell 7, 417-429.

Jefferson et al., (1987) EMBO J. 6, 3901-3907.

Koβmann et al.,(1995) WO 95/04826.

Koβmann et al., (1996) WO 96/19581.

Kortsee *et al.*, (1996) The Plant Journal 10, 83-90.

Logemann *et al.*, (1987) Anal. Biochem. 163, 21-26.

Molulle *et al.*, (1996) Plant Cell 8,1353-1366.

Murray & Thompson (1980) Nucl. Acids Res. 8, 4321- 4325.

Nakamura (1996) Plant Science 121, 1-18.

Nakamura *et al.*, (1996) Planta 199, 209-218.

Pan & Nelson (1984) Plant Physiol. 74, 324-328.

Renz *et al.*, (1995) Genbank accession number X83969.

Sambrook *et al.*, (1989) Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Tolmasky & Krisman (1987) Eur. J. Biochem. 168, 393-397.

Takeda *et al.*, (1993) Carbohydrate Research 240, 253-263.

Visser *et al.*, (1989) Plant Sci. 64, 185-192.

26

Claims

1. A nucleic acid sequence obtainable from potato plants and carrying at least a portion of an isoamylase-type debranching enzyme (DBE) gene.

2. A sequence according to claim 1, comprising at least 200bp of an isoamylase gene.

3. A sequence according to claim 1 or 2, comprising at least 300-600bp of an isoamylase gene.

4. A sequence according to any one of claims 1, 2 or 3, comprising substantially any one of the nucleic acid sequences shown in Figures 5A and 5B, or a functional equivalent thereof.

5. A sequence according to any one of claims 1, 2 or 3, comprising substantially any one of the nucleic acid sequences shown in Figures 6A, 6B, 7A and 7B, or a functional equivalent thereof.

6. A sequence according to any one of claims 1, 2 or 3 which, when linked in the sense orientation to a promoter, encodes a polypeptide comprising the amino acid sequence shown in Figure 8.

7. A sequence according to claim 6, comprising substantially nucleotides 79-1595 of the nucleic acid sequence shown in Figure 8, or a functional equivalent thereof.

8. A nucleic acid construct comprising a sequence in accordance with any one of the preceding claims operably linked in the sense or antisense orientation to a promoter active in a plant.

9. A host cell into which has been introduced a sequence in accordance with any one of claims 1-7, or a construct in accordance with claim 8.

27

10.   A plant cell according to claim 9.

11.   A plantlet grown from a plant cell according to claim 10.

12.   A method of altering a plant by inhibiting the expression of an isoamylase therein, the method comprising the steps of: introducing into the plant at least 200bp of a sequence carrying an isoamylase gene and exhibiting at least 80% identity with a nucleic acid sequence shown in any one of Figures 5A, 5B, 6A, 6B, 7A, 7B or 8, said sequence being operably linked in the sense or antisense orientation to a suitable promoter active in the plant;
and causing transcription of the introduced nucleic acid sequence, said transcript and/or the translation product thereof being sufficient to interfere with the expression of a homologous isoamylase gene naturally present in the plant.

13.   A method according to claim 12, wherein the plant is: potato, sweet potato, maize, wheat, barley, oat, cassava, pea, or rice.

14.   A method according to claim 12 or 13, comprising the introduction of one or more additional nucleic acid sequences, operably linked in the sense or anti-sense orientation to a suitable promoter active in the plant, and causing transcription of the one or more additional nucleic acid sequences, said transcript(s) and/or the translation product(s) thereof being sufficient to interfere with the expression of a homologous gene present in the plant.

15.   A method according to claim 14, wherein said additional nucleic acid sequence comprises an effective portion of a pullulanase type debranching enzyme.

16.   A method according to claim 14 or 15, wherein the said additional nucleic acid sequence comprises a portion which confers a waxy (i.e. low amylose starch) phenotype on the plant.

17. A method according to any one of claims 12-15, wherein the sequences are introduced into a plant having a waxy (i.e. low amylose starch) phenotype.

28

18. A method according to any one of claims 12-17, wherein the method results in a plant comprising starch having different properties compared to starch from an equivalent, but unaltered plant.

19. A plant altered by the method of any one of claims 12-18, or the progeny thereof.

20. Starch obtained from a plant according to claim 19, having altered properties compared to starch obtained from an equivalent but unaltered plant.

21. Starch according to claim 20, having increased branching and/or shorter chain length compared to starch extracted from an equivalent but unaltered plant.

22. Starch according to claim 19 or 20, having reduced peak viscosity compared to starch extracted from an equivalent but unaltered plant, as determined by analysis of a 7.4% (w/v) starch suspension using a Rapid Visco Analyser series 4 instrument.

23. Starch according to any one of claims 19, 20 or 21, having higher setback viscosity compared to starch extracted from an equivalent but unaltered plant, as determined by analysis of a 7.4% (w/v) starch suspension using a Rapid Visco Analyser series 4 instrument.

24. Starch according to any one of claims 19-22, having an increased viscosity onset temperature compared to starch extracted from an equivalent but unaltered plant, as determined by analysis of a 7.4% (w/v) starch suspension using a Rapid Visco Analyser series 4 instrument.

25. A nucleic acid sequence substantially as hereinbefore defined and with reference to the accompanying drawings.

26. A method of altering a plant substantially as hereinbefore defined.

Fig.1.

Fig.1 (Cont).

Multiple sequence alignment (residues ~210–330) for A10906_1, A37035, HIU32761_1, ISOA_PSESP, and ZMU18908_1.

Block 1 (positions 210–240):

Start	Sequence	ID
421	VRPAAGGSGLDLFAEPWAIGGNSYQLGGFPQ--GWSEWNG	A10906_1
447	VRPAAGGSGLDLFAEPWAIGGNSYQLGGFPQ--GWSEWNG	A37035
357	NEPSLQ--NIKLIAEPWDIGHYGYQVGNFPSY--FAEWND	HIU32761_1
447	VRPA	ISOA_PSESP
490	NDPILG--NVKLIAEAWDAGGL-YQEGQFPHWNVWSEWNG	ZMU18908_1

Block 2 (positions 250–280):

Start	Sequence	ID
459	LFRDSLRQAQNELGSMTIYVTQDANDFSGSSNLFQSSGRS	A10906_1
485	LFRDSLRQAQNELGSMTIYVTQDANDFSGSSNLFQSSGRS	A37035
393	RFRDDL-----------------------------	HIU32761_1
450	------------------------------------	ISOA_PSESP
527	KYRDTVRQ----FI-KGTDGFAGA-----------	ZMU18908_1

Block 3 (positions 290–320):

Start	Sequence	ID
499	PWNSINFIDVHQGMTLKDVYSCNGANNSQAWPYGPSDGGT	A10906_1
525	PWNSINFIDVHDGMTLKDVYSCNGAN	A37035
399	------------------------CR	HIU32761_1
450	------------------------------	ISOA_PSESP
546	--FAE--------------CL----	ZMU18908_1

Block 4 (position 330):

Start	Sequence	ID
539	STNYSWDQGMSSA	A10906_1
550		A37035
400		HIU32761_1
450		ISOA_PSESP
550		ZMU18908_1

Fig.2A.

```
              ↓10        ↓20        ↓30        ↓40        ↓50        ↓60        ↓70        ↓80
134-5'.seq  CATATGGATTATGGATGTTGTGTATAACCATTACATGCAAGTGTCCCTTAGTGACTTCTCTGTTCTAGACAAGGTGA
135-5'.seq           GGATCCGATTATGGATGTTGTGTACAACCATTACATGCAAGTGTCCTTTAGTGACTTCTCTGTTCTAGACAAGGTGA
                      ↑40        ↑50        ↑60        ↑70        ↑80        ↑90        ↑100       ↑110

              ↓90        ↓100       ↓110       ↓120       ↓130       ↓140       ↓150       ↓160
134-5'.seq  CGAATCATTGTGTTCCACCTCATTAAACTTATGGTTATAGCCTTATAAAACTTGGTCTCCTTATACGTATTTGTAT
135-5'.seq  CAAACTATTGTGTTCCACCTCATCAAACTTTATGGTTATAGCCTTATAAAACTTGTCTCCTTATAC TATTTGTAT
                      ↑120       ↑130       ↑140       ↑150       ↑160       ↑170       ↑180       ↑190

              ↓170       ↓180       ↓190       ↓200       ↓210       ↓220       ↓230       ↓240
134-5'.seq  TGGGTTTTCTTTCAAAAGATCGTTCCAGGTTACTTTCTAAGGAGAAATGCTGATGGTGGCATTGAAAATAGCACATGCGT
135-5'.seq  CGGGTTTTGGTTTCAATAGATCGTTCCAGGTTACTTTCTAAGGAGAAATGCTGATGGTGGCATTGAAAATAGCACATG GT
                      ↑200       ↑210       ↑220       ↑230       ↑240       ↑250       ↑260       ↑270

              ↓250       ↓260       ↓270       ↓280       ↓290       ↓300       ↓310       ↓320
134-5'.seq  AAACAATACCGCCAGTGAGCATTCATGGTTGAGCGGCTTAATTCTTGACGACCTCAAGTGTCTGGGCAGTTCATTATAAGG
135-5'.seq  AAGCAATACCGCCAGTGAGCATTCATGGTTGAGCGGCTTAATTCTTGACGACCTCAAGTGTCTGGGCAGTTCATTATAAGG
                      ↑280       ↑290       ↑300       ↑310       ↑320       ↑330       ↑340       ↑350

              ↓330       ↓340       ↓350       ↓360       ↓370       ↓380       ↓390       ↓400
134-5'.seq  TATGCTTCCCTCCACCCTCGAATGCCATTGCTACGATATGCTACGATAGGGAATTGTAGAACGTAAATTTGTTCTTCCGTTTTAAGGT
135-5'.seq  TATGCTTCCCTCCACCCTCGAA GCCATTGCTACGATATGCTACGATAGGGAATTGTAGAAC AAAT TTC TC GTTTAAGGT
                      ↑360       ↑370       ↑380       ↑390       ↑400       ↑410

              ↓410       ↓420       ↓430       ↓440       ↓450       ↓460
134-5'.seq  TTCCTCCGTTTTGTTGGCCGAATATCTGTTAACGCTATGGGCTGCGTT-CTCATATTTAAGGACA
135-5'.seq  T CCTCC GTTT TTGGC GAATATCTGTTAACGC GGCTGCGTT CTCATATTT AGGACA
                      ↑440       ↑450       ↑460       ↑470       ↑480       ↑490       ↑500
```

Fig.2B.

```
              ↓40        ↓50        ↓60        ↓70        ↓80        ↓90       ↓100       ↓110
134-3'.seq   GGGATCCGATTTCAAAACGGAAGCCATCAATCTGCAGAAGACGAACAAGGTTAACACATATATTCTCTTTCAGAGAAGA-
                                                                                            GATTTCAAAACGGAAGCCATCAATCTGCA AAGACGAACAAGGTTAACACATA ATTCTCTTTCAGAGAAGA
135-3'.seq   TCATATGGATTTCAAAACGGAAGCCATCAATCTGCAAAAGACGAACAAGGTTAACACATCACATTCTCTTTCAGAGAAGAA
                        ↑10        ↑20        ↑30        ↑40        ↑50        ↑60        ↑70       ↑80

                        ↓120       ↓130       ↓140       ↓150       ↓160       ↓170       ↓180       ↓190
134-3'.seq   TAGGGATTCTTACTCAAGGACAAAAGCATCAGCGTGTACACACATAGAGAACACATAGTGTGACATCAATTGT
             TAGG ATTCTTACTCAAGGACAAAAGCAT AGCGTGTACACACATAGAGAACACATAGTGTGACATCAATTGT
135-3'.seq   TAGGAATTCTTACTCAAGGACAAAAGCATAAGCGTGTAC ACACATAGAGAACACGAAATGGGATATGTGACATCAATTGT
                        ↑90        ↑100       ↑110       ↑120       ↑130       ↑140       ↑150      ↑160

                        ↓200       ↓210       ↓220       ↓230       ↓240       ↓250       ↓260       ↓270
134-3'.seq   GTGAATTAAGTAAGATAAAGTTAACTGATTTTATTAGTTGGTTTGACATATAATAGAGTAAACATTAACTTCT
             GTGAATTAAG A TA           C        TT         TAGGGT A            AGAG   TT C TC
135-3'.seq   GTGAATTAAGCAGTGTATCAAAGGCGAGGCTCAAACATGTTGAGAGCTTCGCTTCGCCTCG
                        ↑170       ↑180       ↑190       ↑200       ↑210       ↑220       ↑230      ↑240

                        ↓280       ↓290       ↓300       ↓310       ↓320       ↓330       ↓340       ↓350
134-3'.seq   GTGTTTCTTGAGAAATCACTTAGTAACTGTTTAGGCATTGGCCATGTTCAGTCATCAACATCAAGAACATAAAGCATG
             TTTT             A   T       GTT      A    C T    G CACA       A    A   AAG
135-3'.seq   CTTTATGTGTGCTTCAGTTGTTGTCATCAAGGTTCTAAGACAAACTTTTCCTTGCCAACGAGCCTCTTATGAAGAAGTGAC
                        ↑250       ↑260       ↑270       ↑280       ↑290       ↑300       ↑310      ↑320

                        ↓360       ↓370       ↓380       ↓390       ↓400       ↓410       ↓420       ↓430
134-3'.seq   GTTCGGTGCGATATTTGAGCAATAAGAACTGTTTATGGGACTGTTAAAGCAATGAACTGCCCATCTAAAGATCAAAC
             T             TT   C TA A      TTT T        TT                     T A  T  AA
135-3'.seq   ACTAAACAATTGATATTTCACTTTATCATAGATTTTTTTCAATTCTTTGGTCATATATTTGTCATTCATGTTTATAAT
                        ↑330       ↑340       ↑350       ↑360       ↑370       ↑380       ↑390      ↑400

                        ↓440       ↓450       ↓460       ↓470       ↓480       ↓490       ↓500
134-3'.seq   TGTTAGAGAAGGGACACTTTTATTGATTTACTTCTAGGTCCTCCATACGTCCCGTGAAATAGTTAT
             T TTA      GG   A   TAT   ATTT   T CTCC   G  CC T   A
135-3'.seq   TATTAATCTTGGACTAAAAAATATACATTTGTATTTTTTTCTCCCCCTTGCGCCCTTTTTTATTAAAG
                        ↑410       ↑420       ↑430       ↑440       ↑450       ↑460
```

Fig.3A.

```
                                                                         Xba I
                                                                          ....
CATATGGATTATGTTGTGTATAACCATTTACATGCAAGTGGTCCCTTTAGTGACTTCTGTGTTCTAGACAAGGTGACGAATCATTGTTCCACCTCATTAAACTTTATGGTTATA
                                                                                                                  + 120
GTATACCTAATACTAATAAATGGTAAATGTACGTTCACCAGGGAAATCACTGAAGAGACAAGATCTGTTCCACTGTTAGTAACACAAGTGGAGTAATTTGAAATACCAATAT

         M  D  V  V  Y  N  H  L  H  A  S  G  P  F  S  D  F  S  V  L  Q  K
                                         ──────exon──────                        ──────intron──────

         SnaB I
         ....
GCCTTATAATAAACTTGGTCTCCTTATACGTATTTTGTATTTGGGTTTCTTTCAAAAGATCGTTCCAGGTTACTTTCTAAGGAGAAATGCTGATGGTGGCATTGAAAATAGCACATGCGT
                                                                                                              + 240
CGGAATATTATTTGAACCAGAGGAATATGCATAAAACATACAAAGAAAGTTTTCTAGCAAGGTCCAATGAAAGATTCCTCTTACGACTACCACCGTAACTTTATCGTGTACGCA

         ──────intron──────                              I  V  P  G  Y  F  L  R  R  N  A  D  G  G  I  E  N  S  T  C  V
                                                        ──────────────────exon──────────────────

AAACAATACCGCCAGTGAGCATTTCATGGTTGAGCGCTTAATTCTTGACGACCTCAAGTGCTGGGCAGTTCATTATAAGGTATGCTTCCTCCACCCTCGAATGCCATTGCTACGATATG
                                                                                                              + 360
TTTGTTATGGCGGTCACTCGTAAAGTACCAACTCGCGAATTAAGAACTGCTGCTGGAGTTCACGACCCGTCAAGTAATATTCCATACGAAGGAGGTGGGAGCTTACGGTAACGATGCTATAC

         N  N  T  A  S  E  H  F  M  V  E  R  L  I  L  D  D  L  K  C  W  A  V  H  Y  K
         ────────────────────────exon────────────────────────               ──────intron──────

GGAATTGTAGAACGTAAATTTGTTTCTCCGTTTGTTGTTGGCCGAATATCGTTTAACGCTATGGGCTGCGTTCTCATATTTTAAGGACATGTGAATGTTTG
                                                                                              + 480
CCTTAACATCTTGCATTTAAACAAAGAGGCAAATTCCAAAGGAGGCAAAATTGCGATACCCGACGCAAGAGTATAAAATTCCTGTACACTTACAAAC

         ──────────────────────────intron──────────────────────────

TATTTTATATAACCCCACC
                   ↑ 500
ATAAAATATTATTGGGGTGG

         ──────intron──────
```

Fig.3B.

```
ATAACTATTTCACGGGACGTATGGAGGACCTAGAGTAAATCAATAAAAGTGTCCCTTCTCTAACAGTTGATCTTTAGATGGGCCAGTTCCCATAAACAGTTCT   381
TATTGATAAAGTGCCCTGCATACCTCCTGGATCTCATTTAGTTATTTCACAGGGATTGTCAAACTAGAAAATCTACCCGTCAAGTAACGAAATTGTCAGGGTATTTGTCAAGA
                                                                                              ┌── intron
                                                                                              └── intron

TATTGCTCAAATATCGCACCGAAACCATGCTTTATGTTCTTGATGTTGATGACTGAAACATGCCAAATGCCTAAAACAGTTACTAAGTGATTTCTCAAGAAACACAGAAGTTAATGTTTA   261
ATAACGAGTTTATAGCGTGGCTTTGGTACGAAATACAAGAACTACAACTACTGACTTGTACCGTTACGGATTTGTCAATGATTCACTAAAAGAGTTCTTTGTGTCTTCAATTACAAAT
                                                                                              ┌── intron
                                                                                              └── intron

                Hinc II
CTCTATTATATGTCAAACCAACTCATAACTAATAAAATCAGTTAACTTTATCTTACTTAATTCACACAATTGATGTCACATATCCCATTCGTGTTCTCTATGTGTGTACACGCTGATGC   141
GAGATAATATACAGTTTGGTTGAGTATTGATTATTTTAGTCAATTGAAATAGAATGAATTAAGTGTGTTAACTACAGTGTATAGGGTAAAGCACAAGAGATACACACATGCGGACTACG
                                                                                              ┌── intron
                                                                                              └── intron

                                                              BamHI  XmaI SmaI KpnI SacI
          Pst I        Hinc II
TTTTGTCCTTGAGTAAGAATCCCTATCTTCTCTGAAAGAGAATATATGTGTTAACCTTGTTCGTCTTCTGCAGATTGATGGCTTCCGTTTTGAAATCGGATCCCCGGGTACCGAGCTCG   21
AAAACAGGAACTCATTCTTAGGGATAGAAGAGACTTTCTTATATACACAATTGGAACAAGCAGAAGACGTCTAACTACCGAAGGCAAAACTTTAGCCTAGGGGCCCATGGCTCGAGC
                                                                                     I  D  G  F  R  F
          ┌── intron                                                                    └────── exon
          └── intron

AATTGCACTGCCCTCGTTT   1
TTAACGTGACCGGCAGCAAA
```

Fig.4A.

```
          EcoRI    SacI        KpnI
                              XmaI
                               SmaI      BamHI
AAAACGACGGCCAGTGAATTCGAGCTCGGTACCCGGGGATCCGATTATGGATGTTGTGTACAACCATTTACATGCAAGTGGTCCTTTTAGTGACTTCTCTGTTCTGTTAGACACAAGGTGACAA    120
TTTTGCTGCCGGTCACTTAAGCTCGAGCCATGGGCCCCTAGGCTAATACCTACAACACATGTTGGTAAATGTACGTTCACCAGGAAAATCACTGAAGAGACAAGATCTGTTCCACTGTT

                                        D  V  V  Y  N  H  L  H  A  S  G  P  F  S  D  F  S  V  L  Q  K
                                                          exon                                        intron

ACTATTGTGTTCCACCTCATCAAACTTTATGGTTATAGCCTTATAATAAACTTGATCTCCTTATACATATTTGTATCGGGTTTGGTTCAATAGATCGTTCCAGGTTACTTTCTAAGGA
TGATAACAACAAGGTGGAGTAGTTTGAAATACCAATATCGGAATATTATTGAACTAGAGGAATATGTATAAAACATAGCCAAACAAGTTATCTAGCAAGGTCCAATGAAAGATTCCT    240
                         intron                                                                          I  V  P  G  Y  F  L  R
                                                                                                                  exon

GAAATGCTGATGGTGGCATTGAAAAATAGCACACATGTGTAAGCAATACCGCCAGTGAGCAATTCATGGTTGAGCGCTTAATTCTTGACGACCTCAGTGCTGGGCAGTTCATTATAAGGTAT
CTTTACGACTACCACCGTAACTTTATCGTGTACACATTCGTTATGGCGGTCACTCGTAAAGTACCAACTCGGAGTTCGAATTAAGAACTGCTGGAGTCACGACCCGTCAAGTAATATTCCATA    360

         R  N  A  D  G  G  I  E  N  S  T  C  V  S  N  T  A  S  E  H  F  M  V  E  R  L  I  L  D  D  L  K  C  W  A  V  H  Y  K  [intr]
                                  exon                                                                                    [exon]

GCTTCCCTCCACCCTCGAACGCCATTGCTACGATATGGGAATTGTAGAACCCAAATCCCTTCCTCGTTTAAGGTTCCTCCTGTTTAATTGGCTGAATATCGTGTTTAACGCATAGGCTG
CGAAGGGAGGTGGGAGCTTGCGGTAACGTACACCCTTAACACTCTTGGGTTTAGGGAAGGAGGACAAAATTCCAAGGAGGACAAAATTAACCGACTTATAGACAAATTGCGTATCCGAC    480
                                                  intron

CGTTACTCATATTTAGGACA
GCAATGAGTATAAATCCTGT
              500
          intron
```

Fig.4B.

GGTGGGGGATTAAGGCGCAAAATGTAAATAAGGCGTGGGCTTTAATAAAAAGGGGCGCAAGGGGAGAAAAATACAAATGTATATTTTAGTCCAAGATTAATAATATAAACATGAATG 385
CCACCCCCTAATTCCGCGTTTTATTCGCACTTTATTTCCCGCGTTTCCCCTCTTTTTTATGTTTACATATAAAAATCAGGTTCTAATTATTAATATTTGTACTTAC

intron

ACAAATATGACCAAAGAAATTGAAAAAAATCTATGAAATGAAATCAATTGTTAGTGTCACTTCTTCATAAGAGGCTCGTTGGCAAGGAAAGTTTGTCTTAGAACCTTGAT 28
TGTTTATATACTGGTTTCTTCTTAACTTTTTTTTAGATACTATTCACTTTATAGTTAACAAATCACAGTGAAGAAGTATTCCGAGCAACCGTTCCTTTCAAACAGAATCTTGGAACTA

intron

BssH II

GACAACACTGAAGCACACATAAAGCGAGGCGAAGCTCTCAACATGTGTTTGAGGCGCATAAGGCGCCTTTGATACACACAATTGATGTCACAT 145
CTGTTGTGACTTCGTCGTGTGTATTTCGCTCCGCTTCGAGAGTTGTACAAACTCGAGCGCGGAAATCCCGTATTCGCGCCGGAAACTATGTGACGAATTAAGTGTGTTAACTACAGTGTA

intron

EcoR I

ATCCCATTTCGTGTTCTCTATGTGTGTACACGCTTATGCTTTGTCCTTGAGTAAGAATTCCTATTCTTCTCTGAAAGAGAATGTATGTGTTAACCTTGTTCGTCTTTTTGCAGATTGATG 25
TAGGGTAAAGCCACAAGAGATACACACATGTGCGAATACGAAACAGGAACTCATTCTTAAGGATAAGAGAAGACTTTCTTACATACACAATTGGAACAAGCAGAAAACGTCTAACTAC

Hinc II

intron

I D
└─ exon ─

GCTTCCGTTTGAAATCCATATGA
CGAAGGCAAAACTTTAGGTATACT 1

G F R F
└─ exon ─

Fig.5A.

```
                    ↓10       ↓20       ↓30       ↓40       ↓50       ↓60       ↓70       ↓80
133-5'.seq   CATATGGATTATGGATGTTGTGTTTAACCACACTGCTGAAGGAAATGGTCCCATACTATCATTTAGAGGCATTG
                          GATTATGGATGTTGTGT
137-5'.seq   GGATCCGATTATGGTGTGTGTACAATCACACTGCTGAAGGAAATGGTCCCATACTATCATTTAGAGGCATTG
                  ↓40       ↓50       ↓60       ↓70       ↓80       ↓90       ↓100      ↓110

133-5'.seq   ACAACAGTGTGTTTTATACGCTCCTAGCTCCTAAGGTATGTTGTCGCTAATCCATGCACATTTTTGACTCTTAAA
137-5'.seq   ACAACAGTGTGTTTTATACGCTCCTAGCTCCTAAGGTATGTTGTCGCTAATCCATGCACATTTTTGACTCTTAAA
                  ↓90       ↓100      ↓110      ↓120      ↓130      ↓140      ↓150      ↓160
             ACAACAGTGTGTTTTATACGCTCCTAGCTCCTAAGGTATGTTGTCGCTAATCCATGCACATTTAGTTTTGACTCTTAAA
                          ↓120      ↓130      ↓140      ↓150      ↓160      ↓170      ↓180      ↓190

133-5'.seq   ATAAAGTTCAGAGTTCCATATAGGATTCATTGCTTTCTTCTTCAGAATTAATCATTGTGAAGTTGC
137-5'.seq   ATAAAGTTCAGAGTTCTAAGTTGAAGTTCCATATATGGATTCATTGCTTTCTTCTTCAG ATTTAATCATTGTGAAGTTGC
                  ↓170      ↓180      ↓190      ↓200      ↓210      ↓220      ↓230      ↓240
             ATAAAGTTCAGAGTTCTAAGTTGAAGTTCCATATATGGATTCATTGCTTTCTTCTTCAGATTTAATCATTGTGAAGTTGC
                          ↓200      ↓210      ↓220      ↓230      ↓240      ↓250      ↓260      ↓270

133-5'.seq   TCCACTATTACAAATTGATTCTTCGGATCATGAAAATATCCTTTTT--ATTCTAAGATAATTAAACCCTGCAATTGCT
137-5'.seq   TCCACTATTACAAATTGATTCTTCGGA CATGA CATGAAAATATCCTTTTTT ATTCTAAGATAATTAAACCCTGCAATTGCT
                  ↓250      ↓260      ↓270      ↓280      ↓290      ↓300      ↓310
             TCCACTATTACAAATTGATTCTTCGGAACATGAAAATATCCTTTTTTTCTATTCTAAGATAATTAAACCCTGCAATTGCT
                          ↓280      ↓290      ↓300      ↓310      ↓320      ↓330      ↓340      ↓350

133-5'.seq   TTACTTATAAACAAAAAAAAAAAATAATTAAACCCTGCGGGTGTGGG-ATTCTCCACCTTAAGCTTATTATGTTGTCTATAAC
137-5'.seq   TTACTTATAAACAAAAAAAAAAATAATTAAACCCTGC GTGTGGG ATTCTCCACCTTAAGC TATTATGTTGTCTATAAC
                  ↓320      ↓330      ↓340      ↓350      ↓360      ↓370      ↓380      ↓390
             TTACTTATAAACAAAAAAAAAATAATTAAACCCTGCGGTGTGGGGATTCTCCACCTTAAGCCTATTATGTTGTCTATAAC
                          ↓360      ↓370      ↓380      ↓390      ↓400      ↓410      ↓420      ↓430

133-5'.seq   TCTATATC--ACACACTTCATCCAAATCATAACAGATGAACTGGTTATATCCTCTCCAAAAAATCATA
137-5'.seq   TCTATATC ACACACTTC ACACACTTCCTCCCAAATCATAACAGATGAACTGGT ATATCCTC AAAAA A
                  ↓400      ↓410      ↓420      ↓430      ↓440      ↓450      ↓460      ↓470
             TCTATATCTCACACACTTCCTCCCAAATCATAACAGATGAACTGGTCATATCCTCTCCAAAAAATCATAA
                          ↓440      ↓450      ↓460      ↓470      ↓480      ↓490      ↓500
```

Fig.5B.

```
                ↓40        ↓50        ↓60        ↓70        ↓80        ↓90        ↓100       ↓110
133-3'.seq   GGATCCGATTTCAAAACGGAAGCCATCTACGTGCATTTCGGTAACCCAATATCTGTCAAGAGATCAAAGTAGACCTTTTC
137-3'.seq      GATTTCAAAACGGAAGCCATCTACGTGCATTTCGGTAACCCAATATCTGTCAAGA ATCAAAG AGACCTTTTC
             CATATGGATTTCAAAAC GGAAGCCATCTACGTGCATTTCGGTAACCCAATATCTGTCAAGAAAT CAAAGCAGACCTTTTC
                           ↑10        ↑20        ↑30        ↑40        ↑50        ↑60        ↑70    ↑80

                ↓120       ↓130       ↓140       ↓150       ↓160       ↓170       ↓180       ↓190
133-3'.seq   TATTATTATCAATATAGGACATTTCATGGTTAGAGCGGCCAAATTTGTAACTCATCATGCAAGATATAGACTTTATAGGGAA
137-3'.seq   TATTATTATCAATATAGGACATTTCATGGTTAGAGCGGCCAAATTTGT ACTCATATGCA GATATAGACTTTATAGGGAA
             TATTATTATCAATATAGGACATTTCATGGTTAGAGCGGCCAAATTTGTGACTCATGCAGGATATAGACTTTATAGGGAA
                           ↑90        ↑100       ↑110       ↑120       ↑130       ↑140       ↑150    ↑160

                ↓200       ↓210       ↓220       ↓230       ↓240       ↓250       ↓260       ↓270
133-3'.seq   CTCCATGCATTAAAACCACAGAGAAATTTGAGCCTTCAAAATAGCAGAGAAGTTCATCTAGAAAATCCCACTTAGACTACAA
137-3'.seq   CTCCATGCATTAAAACCACAGAGAAATTTGAGCCTTCAAAATAGCAGAGAAGTTCATCTAGAAAATCCCACTTAGACTACAA
             CTCCATGCATTAAAACCACAGAGAAATTTGAGCCTTCAAAATAGCAGAGAAGTTCATCTAGAAAATCCCACTTAGACTACAA
                           ↑170       ↑180       ↑190       ↑200       ↑210       ↑220       ↑230    ↑240

                ↓280       ↓290       ↓300       ↓310       ↓320       ↓330       ↓340       ↓350
133-3'.seq   AATTAGACTGAAGATTAGATGCAGACACAAGATAAAACAATTGATACACACTGAGAACTAACACAATAATATTCTTTAGTC
137-3'.seq   AATTAGACTGAAGATTAGATGCAGACACAAGATAAAACAATTGATACACACTGAGAACTAACACAATAATATTCTTTAGTC
             AATTAGACTGAAGATTAGATGCAGACACAAGATAAAACAATTGATACACACTGAGAACTAACACAATAATATTCTTTAGTC
                           ↑250       ↑260       ↑270       ↑280       ↑290       ↑300       ↑310    ↑320

                ↓360       ↓370       ↓380       ↓390       ↓400       ↓410       ↓420       ↓430
133-3'.seq   CACTTGTGAAAAGCTTAAGGGAAACCTCAGCTATAAATGCTTTAAAAGTCGACGTTTCTTTAAACCAAAAGCAGAAAAAC
137-3'.seq   ACTTGTGAAAAGCTTAAGG AAACCTCAGCTA  AAATGCTTTAAAAGTCGACGTTTCTTTAAAC AAAAGC AAAAC
             AACTTGTGAAAAGCTTAAGG-AAACCTCAGCTACAAATGCTTTAAAAGTCGACGTTTCTTTAAACAAAAGCCAAAAAC
                           ↑330       ↑340       ↑350       ↑360       ↑370       ↑380       ↑390

                ↓440       ↓450       ↓460       ↓470       ↓480       ↓490       ↓500
133-3'.seq   ATAGGCACTGAAACCCACTTAATTATTTATTTTTGTAGCTTCATTCCCTCATTCCCTCATAGACATGAAATG
137-3'.seq   ATAGGCACTG AACCCACTTAATTATTTATTTTTGTAGCTTCATTCCCTCATTCCCTCATAGACATGAAATG
             ATAGGCACTGGAACCCACTTAATTATTTATTTTTGTAGCTTCATTCCCTCATTCCCTCATAGACATGAAATG
                           ↑400       ↑410       ↑420       ↑430       ↑440       ↑450       ↑460
```

Fig.6A.

Nde I

```
                                                                              120
CATATGGATTATGGATGTGTTAACCACACTGCTGAAGGAAATGAAAATGGTCCATACTATCATTTAGAGGCATTGACAACAGTGTGTTTATACGCTAGCTCCTAAGGTATGTTT
GTATACCTAATACCTACACACAAATTGGTGTGACGACTTCCTTACTTTACTTTACCAGGGTATGGATGTGTCACAAATATGCGATCGAGGATTCCATACAAA

        M  D  V  V  F  N  H  T  A  E  G  N  E  N  G  P  I  L  S  F  R  G  I  D  N  S  V  F  Y  T  L  A  P  K    |intron|
                                        exon
```

```
                                                                              240
GTCGCTAATCCATGCACATTTAGTTGTTTTTGACTCTTAAAAATAAAGTTCAGAGTCTAAGTTGAAGTTCCATATATGGGATTCATTTGCTTTCTCTTCAGAATTTAATCATTGTGAAGTTGC
CAGCGATTAGGTACGTGTAAATCAAAAACTGAGAATTTTATTCAAGTCTCAGATTCAACTTCAAGGTATATACCTAAGTAAACGAAAGAAGAGAAGTCTTAAATTAGTAACACTTCAACG

                                    intron
```

```
                                                                              360
TCCACTATTACAAATTGATTCTTCGGATCATGAAAATATCCTTTTTTTTTATTCTAAGATAATTAAACCCTGCAATTGCTTTACTTATAAACAAAAGAAATAATAATTAAACCCTGCGGTGTG
AGGTGATAATGTTTAACTAAGAAGCCTAGTACTTTTATAGGAAAAAAAATAAGAGATTCTATTAATTAATTTGTTTTTATTATTTGGGACGCCACAC

                                    intron
```

Hind III

```
                                                                              480
GGATTCTCCACCTTAAGCTTATTATGTTGTCTATAACTCTATATCACACACTTCATCCAAATCATAACAGATGAACTGGTTATATCCTCTCCAAAAATCATACAGATGAAATGGTTAGT
CCTAAGAGGTGGAATTCGAATAATACAACAGATATTGAGATAGTGTGTGAAGTAGGTTTAGTATTGTCTACTTGACCAATATAGGAGAGGTTTTTAGTATGTCTACTTTACCAATCA

                                    intron
```

```
                                                                   501
TGTAAGGCTCATNTCCTTATG
ACATTCCGAGTANAGGAATAC

    intron
```

Fig.6B.

Sac I
EcoR I
CGAATTCACTGGCCGTCGTTT 1
GCTTAAGTGACCGGCAGCAAA

intron

BamH I Xma I Sma I Kpn I
GAAATGTCCTATATTGATAATAATAGAAAAGGTCTACTTTGATCTCTTGACAGATATTGGGTTACCGAAATGCACGTAGATGGCTTCCGTTTGAAATCGGATCCCCGGGTACCGAGCT 22
CTTTACAGGATAATAACTATTATTATCTTTTCCAGATGAAACTAGAGAACTTGTCTATAACCCAATGGCTTTACGTGCATCTACCGAAGGCAAAACTTTAGCCTAGGGGGCCCATGGCTCGA

exon

Y W V T E M H H V D G F F R F E

intron

Xba I Nsi I Nde I
GATTTTCTAGATGAACTTCTCTGCTATTTGAAGGCTCAAATTTCTGTGGTTTAATGCATGGAGTTCCCTATAAAGTCTATATCTTGCATATGAGTTACAAATTTGGGCGCTCTAACCAT 142
CTAAAAGATCTACTTGAAGAGACGATAAAACTTCCGAGTTAAAGACACCAAAATTACGTACCTCAAGGGATATTTCAGATATAGAACGTACTCAATGTTTAAACCGCGAGATTGGTA

intron

Hind III Ssp I
GTTTCCCTTAAGCTTTTCACAAGTGGACTAAAGAATATTGTGTTAGTTCTTCAGTGTGTATCAATTGTTTATCTTGTGTCTCGCATCTAATCTTCAGTGTCTAATTTTTGTAGTCTAAGTGG 262
CAAAGGGAATTCGAAAGTGTTCACCTGATTTCTTATAACAACAATCAAGAAGTCACACATAGTTAACAAAATAGAACACAGACGTAGATTAAAACATCAGATTCACC

intron

Sal I Hinc II
CATTTCATGTCTATGAGGGAAATGAAGCTACAAATAAAATAAATTAAGTGGGTTTCAGTGCCTATGTTTTCTGCTTTTGGTTTAAAGAAAAACGTGGACTTTTAAAGCATTTATAGCTGAG 382
GTAAAGTACAGATACTCCCTTACTTCGATGTTTATTTTATTAATTCACCCAAGTCACGGATACAAAGACGAAACCAAATTTCTTTGCAGCTGAAAATTCGTAAATATCGACTC

intron

Fig.7A.

EcoRI   SacI   KpnI / XmaI / SmaI / BamHI

```
AAACGACGGCCAGTGAATTCGAGCTCGGTACCCGGGGATCCGATTATGGATGTGTGTACAATCACACTGCTGAAGGAAATGAAAATGGTCCATACTATCATTTAGAGGCATTGACAAC
TTTGCTGCCGGTCACTTAAGCTCGAGCCATGGGCCCCTAGGCTAATACCTACAACACATGTTAGTGTGACGACTTCCTTACTTTACCAGGGTATGATAGTAAATCTCCGTAACTGTTG
```
120

```
                                  M  D  V  V  Y  N  H  T  A  E  G  N  E  N  G  P  I  L  S  F  R  G  I  D  N
                                                                              exon
```

```
AGTGTGTTTATACGGCTAGCTCCTAAGGTATGTGTTTCGCTAATCCATGCACATTTAGTTTTTGACTCTTAAAAATAAAGTTCAGAGTCTAAGTTGAAGTCTTCCATATATGGATTCATTT
TCACACAAAATGCGATCGAGGAGTTCCATACAAACGATTAGGTACGTGTAAATCAAAAACTGAGAATTTATTTCAAGTCTCAGAGTCTCAAGTTCAAGTATATACCTAAGTAA
```
240

```
S  V  F  Y  T  L  A  P  K
         exon
```

```
GCTTTCTCTTCCAGTATTAATCATTGTGAAGTTGCTCCACTATTACAAATTGATTCTTCGGAACATGAAAATATCCTTTTTCTATTCTAAGATAATTAAACCCTGCAATTGCTTTACT
CGAAAGAGAAGTCATAATTAGTAACACTTCAAGAGGTGATAATGTTTACTTTGTACTTTATAGGAAAAAGATAAGATTCTATTAATTGGGGACGTTAACGAAATGA
```
360

```
                                                    intron
```

PstI

```
TATAAACAAAAAAAATAATAATTAAACCCTGCAGTGTGGGGATTCTCCACCTTAAGCCTATTAATGTTGTCTATAACTCTATATCTCACACACTTCCTCCAAATCATAACAGATGAACTGGT
ATATTTGTTTTGTTTTTATTAATTTGGGACGTCACACCCCTAAGAGGTGGAATTCGGATAATAACAACAGATATTGAGATATAGAGAGTGTGTGAAGGAGGTTTAGTATTGTCTACTTGACCA
```
480

```
                                                    intron
```

```
CATATCCTCTCAAAAAATCATAA
GTATAGGAGAGTTTTTTAGTATT
```
503

```
                                                    intron →
```

Fig.7B.

Fig.8.

Restriction sites indicated: EcoRI, SacI, KpnI, XmaI, SmaI, BamHI, BclI, HindIII, StuI, NcoI, XmnI.

Position markers: 140, 280, 420, 560, 700, 840.

```
GTAAGCCAGGTTTTCCCAGTCACCACGTGTAAAACGACGGCCAGTCGCCGGGATCCGATTGAGCTCGATTGACAACAGTGTGTTTATACCGTCAGTCTCCTAAGGTGAATTTACAACTACTCA
CATTCGGTCCAAAAGGGTCAGTGCTGCTAAACATTTGCTGCCGGTCAGCTCGAGCTAACTCGAGCCGCCCTAGGCTAACTCCGTAACTGTTGTCACCACAAAATGCCAGTCCCACTTAAAATGTTGATGAGT
                                                                                              G I Q N S V F Y T L A P K G E F Y N Y S

GGATGTGGAAATACCTTCAACTGTAATAATCCATTGTACGTCAATTTATAGTGGATTGGTTGAGATATTGGGTACCGAAATGCACGTAGATGGCTTCCCGTTTCATCTTGCTTCTTACAAGGAGTAGCAGCTC
CCTACACCTTTATGGAAGTTGACATTATTAGGGTAACATGCAGTTAAAATATCCTAACCCATGGCTTTACGTGCATCTACCGAAGGCGAAACTAGAACGAAGGATGGAATGTTCTTCATCGTCGAG
                                                                                              G G N T F N C N N P I V R Q F I V D C L R Y V V T E M H V D G F R F D L A S I L T R S S S

GTGGAATGCTCTAAATGTCTATGGAAATTCAATTGACGGTGACGTGATCGATCCACCATTGATTGATAATGATTAGCAATGATCCAATACTTCGTGGAGTAAAGCTTATAGCTGAAGCAT
CACCTTACGACATTTACAGATACCTTAAGTTAACTGCCACTAGCACTGTTCCGTGGTAACTAACTAATCTAATCGTTATGAAGCACCACTCATTTCGAATATCGACTTCCTA
                                                                                              HindIII

GGGATTGTGGAGGCCTTTACCAAGGTTGGCATGTTTCCCGCACTGGGTATCTCGTCGGAGTGCGAACGAACTACGTGACATGGTAGCCAGTTCATCCAAGGCACTGATGGGTTTTCGTGGGGGCTTTGTCGAATGCCTTTGT
CCCTAACACCTCCGGAAATGGTTCAACGGTATTCACAAAGGGGTAGCACTTCCGTGACTACCATGCTGTCAAGTAGGTTCCGTGACTACCATCCGAAAAGACCCGAAAACA
                                                                                              W D C G G L Y Q V G H F P H W G I W S E W N E V R Q D M V R Q F I Q G T D G F S G A F A E C L C

GGAAGCCCAAATCTATACCGAAAAGGAGGAAGAAACCATGCAACAGTATAAAATTCGTGTCGTGCCCACCAGATGCTTTTTACTTTGCTCGATTCATTTAGTACTTGGCTCATTTAGTGACATACAAACACAATTTGCCAATGAGAGGACACAA
CCTTCGGGTTTAGATATGGCTCTTTTCCTCCTTCTTTGGTACCGTTGTCATATTTAAAGCACAGCTGGTCTCTGTAGGGTCATCATTGAAATCTGTATTTGTTATTCGTATTATTCGTTATTGTTTAAACCGTTACCTCGTTATTGTTTAAACCGTTACCTTACCTTCCTTGTGTCTTGTT
                                                                                              G S P N L Y Q K G G R K P W N S I N F V C A H D G F T L A D L V T Y N N K H N L A N G E D N K

AGACGGGGAGAATCACAATAATAGTTGGAATTGCGAATTGCGTGAGGAAGGAGAATTGCAACTACTTCTTATGAGGAAAAAGACAAATGCGGAACTTCTTCCTGCCTTATGGTTTCCCAAGGTGTTCCCATGATAT
TCTGCCCCCTCTTAGTGTTATTATCATCAACCTTAACACCTTAAACGTTCATAGAAATACTTCTTTAACTCCTTTTCTGTTTACGCCTTGAAGAAGGAGAACGGAATACCAAAGGGGTTCCACAAGGGTACTATA
                                                                                              D G E N H N S W N C G E E G E F A S I F M K K L R K R Q H R N F F L C L H V S D G V P H I
```

Fig.8 (Cont).

ATATGGGGATGAATGGTCACACTAAGGGAGGAAACACAACCTATTGCCATGATAATTATTAATTACTTCCGTTGGGATAAGAAGGATGAATCTTCATCTGATTTTTTGAGATTTTGCGGCCTCATGACCAAA
TATACCCGTACTTATCACCAGTGTGATTCCCTCCTTTGTGTTGTGGATAACGGTACTATTAATAATTAATGAAGGCAACCCTATTCTTCCTACTTAGAGTAGTCTAAAAACCTCTAAAAACGCCGGAGTACTGGTTT → 980

Y M G D E Y G H T K G G N N T Y C H D N Y I N Y F R W D K K D E S S S D F L R F C G L M T K

          Sal I
          Hinc II

TTCCGGCCATGAATGTGAATCACTGGGATTAGAGATGGTTTCCCTACAGACCAGAAAGGCTGCAATGGCATGGTCTGAAACAAGTCGATTCGTTGCCATTCACACTGGTCGCGACAAAGTGAA
AAGGCCGGTACTTACACTTAGTGACCCTAATCTCTACCAAAGGGATGTCTGGTCTTTCCGACGTTACCGTACCAGACTTTGTGAGGTCTAACCAGTGTGACCAGCTGTTTCACTT → 1120

F R H E C E S L G L D G F P T A E R L Q W H G H T P R T P D W S E T S R F V A F T L V D K V K

          Stu I

GGCAGAACTATATATTGCCTTAACGCCAGCCATTGCCTGTAACGATTACACTTCCAGATAGGCCTGGTTATAGATGGCAGCCGTTTGTGGACACAGGCAAACCAGCACCATTGACTTCTTGACAGACGACGTTCCTG
CCGTCTTGATATATAACGGAATTGCCGTCGGTAACGTTGCTAATGTGAAGTCTATCCGGACCAATATCTACCGGGGCCAAACACCTGTGTCCGTTTGGTCGTGTAAACTGAAGAACTGTCTGCTGCCAAGGAC → 1260

G E L Y I A F N A S H L P V T I T L P D R P G Y R W Q P F V D T G K P A P F D F L T D D V P

          Ssp I   Nsi I

ACAGAGACACGAGCCAACCAACAATATTCTCATTTTCTGGACGCAACCAGTATCCGATGCTCAGTTATTCATCCATTATTCTTTACTATCATCTGCTGATGATCGTAAGTTCATTCACCAAGTAGTGGGAGGTAAAT
TCTCTCTGTGCTCGGTTGGTTGTTATAAGAGTAAAAGACCTGCGCTTGGTCATAGGCTACGACTCAATAAGTAGTCAATAAGAGAAATGATAGTAGACGACTACTACGTATCAAAGTAAGTGGTTCAATCCACCTCCATTTA → 1400

E R E T A A K Q Y S H F L D A N Q Y P H L S Y S S I I L L S S A D D A

CAGCTTCAGATTTGTTATATGCGTAGGTCGTTACTTTGTAAATAAAGTAGGACAGGACAACGGATAAAATTTGTGAGGAAGAAGCTGATGATTTATAAGATACACCTTGTATTATAATTGC
GTCGAAGTCTAAAACAATATACGCTCACTCCACTGAATGAAACATTTATTTCATTCTTCGATTTAAAGACACTCCTTCTTCGACTGCCTATTTTAAACACTCTCCTTCTTCGACTTGCCTATTTTAAACACTCTCTATGTGGAACATAATAATTAACG

          Cla I   Nde I → 1540

ATTTATATAAAATAAAATACTAGTAGTGAACTTGTCTGTGCCGAAAAAAAAAAAAAAAAAATCCCTATAGTAGTCGTATTATGACTAGTAGATCCTCTAGAGTCGACTGTGG
TAAATATATTTATTTATGTATCACTTGAACAGACACGCTTTTTTTTTTTTTTTTTTTAGGGATATCATCAGCATAATACTGATCATCTAGGAGATCTCAGCTGACACC → 1663

                  Sal I
       Xba I   Hinc II

Fig.9.

Fig.10.

```
                                        Kpn I
                                        Xma I
                                        |  Sma I
      EcoR I         Sac I              |  |  BamH I
AAACGGACGGCCAGTGAATTCGAGCTCGGTACCCGGGGATCCGATTGACGCGATGTCAAGCCTGAGGGCTGGGATAATCTACAAGATGAGAAG
TTTGCTGCCGGTCACTTAAGCTCGAGCCATGGGCCCCTAGGCTAACTGCTACAGTTCGGACTCCCGACCCTATTAGATGTTCTACTCTTC
                                                                                              90

                     D  D  V  K  P  E  G  W  D  N  L  Q  D  E  K

CCAAATCTTCTTTTCTGATGTGTCAGCATCTATGAGCTGCATGTTAGAGATTTCAGGCGCCAATGACCCTACTGTGCCTCATGAATTT
GGTTTAGAAGAAAAGGAAAGACTACAGTCGTAGATACTCGACGTACAATCTCTAAAGTCCGCGGTTACTGGGATGACACGGAGTACTTAAA
                                                                                             180

   P  N  L  L  S  F  S  D  V  S  I  Y  E  L  H  V  R  D  F  S  A  N  D  P  T  V  P  H  E  F
                                             Pvu II

CAGGGCGGGTTATCTCGCCCTTCACGTCGCCAGGACTCAGCTGGTGTCCAACATTTGAAAAGATTATCAAGTGCTGGTATCACTCATGTTCAT
GTCCCGCGCCAATAGAGCGGGAAGTGCAGCGGTCCTGAGTCGACCACAGGTTGTAAACTTTTCTAATAGTTCACGACCATAGTGAGTACAAGTA
                                                                                               270

   Q  G  G  Y  L  A  F  T  S  Q  D  S  A  G  V  Q  H  L  K  R  L  S  S  A  G  I  T  H  V  H

                                                              EcoR V
CTTGCTGCCAACCTATCAATTTGCTGGTGTGTCGAAGATGAGAAGATATCGAGAAACCCAACTCTTTTCCACCA
GACGACGGTTGGATAGTTAAACGACCACAGCTTCTACTCTTCTATAGCTCTTTGGGTTGAGAAAGGTGGT
                                                                         360

   L  L  P  T  Y  Q  F  A  G  V  E  D  E  K  H  K  W  K  Y  T  D  I  E  K  P  N  S  F  P  P
```

GATTCTGAGGAGCAGCAAGCTCTTATCACAGCCATCCAAGATGAAGATGCCTATAAATCCTGGTCTCTGGGGAGTTCCA
CTAAGACTCCTCGTCGTTCGAGAATAGTGTCGGTAGGTTCTACTTCTACGGATATTTAGGACCAGAGACCCCTCAAGGT
D S E E Q Q A L I T A I Q D E D A Y N W G Y N P G L W G V P

450

AAGGGAAGCTACGCTAGTAATGCAAATGGTCCTTGTCGTATCGTTGAGTTTCGGAAAATGGTGCAGGCACTTAATCGTATTGGTCTACGT
TTCCCTTCGATGCGATCATTACGTTTACCAGGAACAGCATAGCAACTCAAAGCCTTTTACCACGTCCGTGAATTAGCATAACCAGATGCA
K G S Y A S N A N G P C R I V E F R K M V Q A L N R I G L R

540

Xba I

GTTGTATTGGATGTTGTTACAATCATTTACATGCAAGTGGTCCTTTTAGTGACTTCTCTGTTCTAGACAAGATCGTTCCAGGTTACTTT
CAACATAACCTACAACAATGTTAGTAAATGTACGTTCACCAGGAAAATCACTGAAGAGACAAGATCTGTTCTAGCAAGGTCCAATGAAA
V V L D V V Y N H L H A S G P F S D F S V L D K I V P G Y F

630

CTAAGGAGAAATGCTGATGGGCATTGAAAATAGCACATGTGTAAACAATACCGCCAGTGAGCATTTCATGGTTGAGCGCTTAATTCTT
GATTCCTCTTTACGACTACCACCGTAACTTTTATCGTGTACACATTTGTTATGGCGGTCACTCGTAAAGTACCAACTCGCGAATTAAGAA
L R R N A D G G I E N S T C V N N T A S E H F M V E R L I L

720

GACGACCTCAAGTGCTGGGCAGTTCATTATAAGATTGATGGCTTCCGATTTGATAATCCATATG
CTGCTGGAGTTCACGACCCGTCAAGTAATATTCTAACTACCGAAGGCTAAACTATTAGGTATAC
D D L K C W A V H Y K I D G F R F D
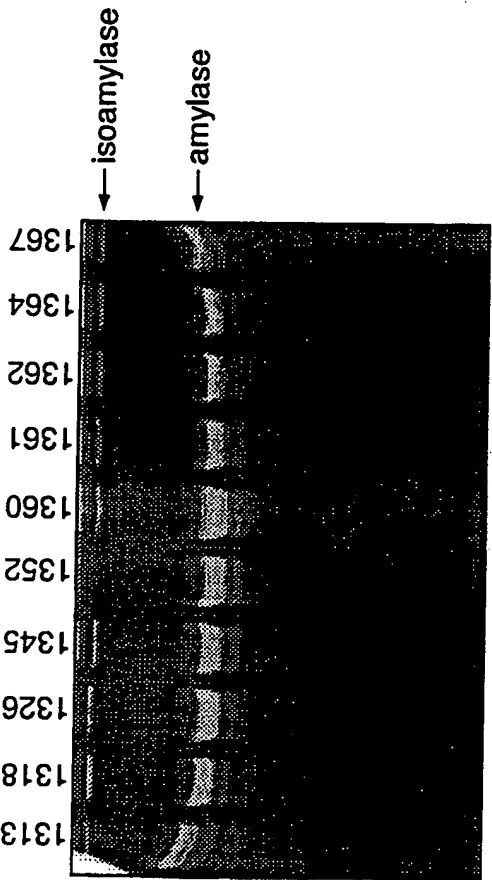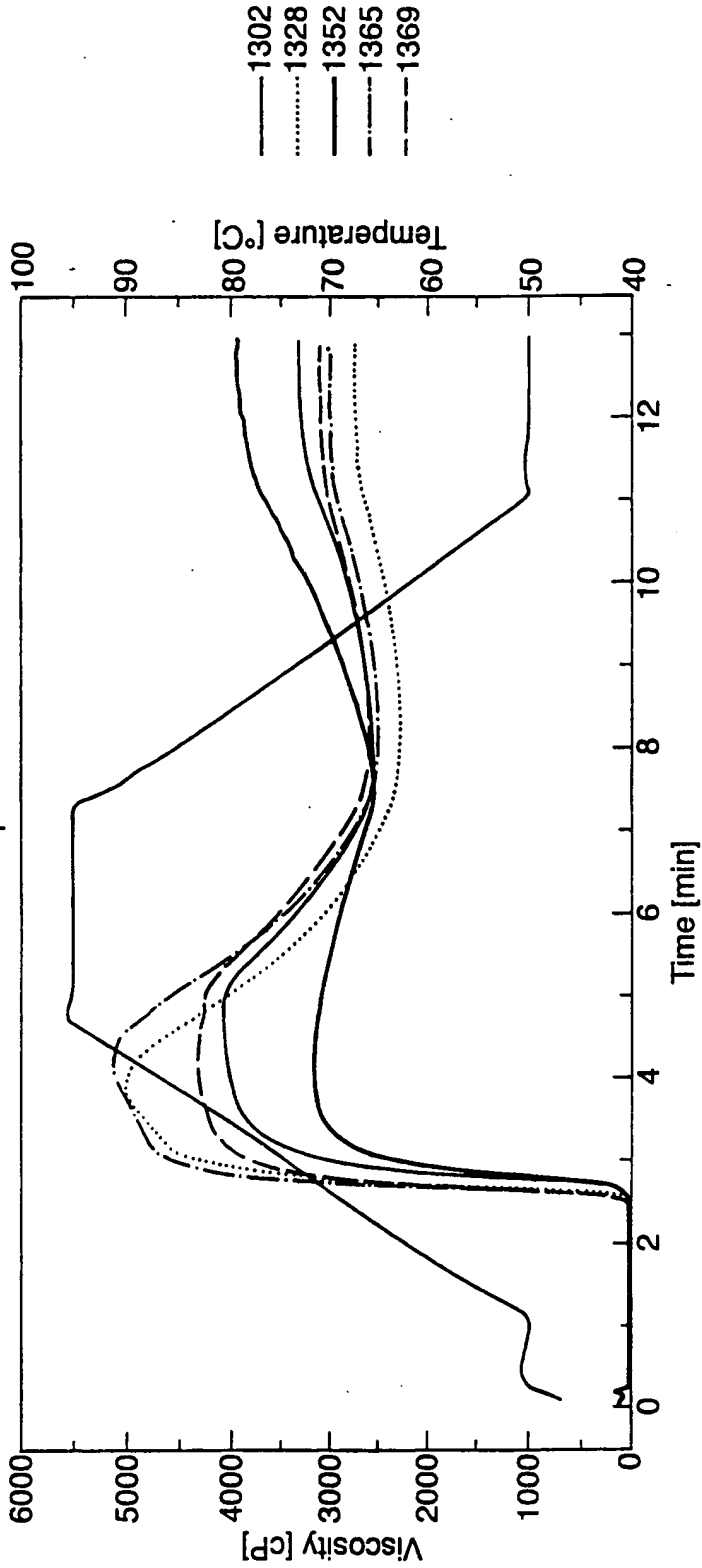
784

Fig.10 (Cont).

Fig.11.
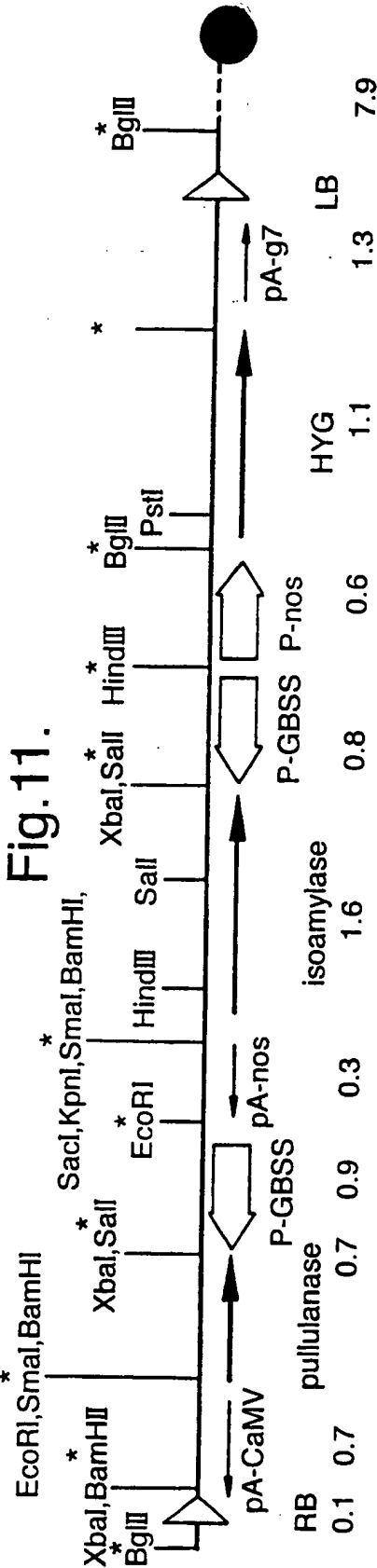
Fig.12.

Fig.13.
RVA profiles

THIS PAGE BLANK (USPTO)